

XX The present sequence represents a peptide linker used in the fusion
CC proteins of the invention. The specification describes a method for
CC improving pathogen resistance or tolerance of a plant. The method
CC comprises transformation of the plant with a transgene encoding a
CC fusion protein which comprises two or more protein or domains that
CC are capable of improving pathogen resistance or tolerance when
CC expressed on their own. Specifically, two distinct proteinase
CC inhibitors are co-delivered, as a fusion, to *Arabidopsis thaliana*. The
CC used to improve pathogen resistance or tolerance of a plant and its
CC descendants, especially against parasitic nematode attack.

Sequence 14 AA;
Query Match
Best Local Similarity 100.0%; Score 70; DB 20; Length 14;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 VILGVGPAKIOPEG 14
1 VILGVGPAKIOPEG 14

RESULT 2
AAB80610
ID AAB80610 standard; Protein; 79 AA.
AC AAB80610;
DT 02-MAY-2001 (first entry)
DE Environmental stress tolerant protein SEQ ID 6.
KW Environmental stress resistance; salt; heat; desert; transgenic plant.
OS *Bruguiera searungula*.
XX WO200106006-A1.
XX 25-JAN-2001.
XX 19-JUL-2000; 2000WO-JP04862.
XX 19-JUL-1999; 99JP-0235910.
XX 24-MAR-2000; 2000JP-0083377.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Yamada A, Ozeki Y, Saito T;
XX WPI: 2001-147355/15.
XX N-PSDB: AAF74189.

Screening method to obtain DNA encoding environmental stress resistance
factor, useful for producing transgenic plants resistant to
environmental stress -
Claim 19; Page 83-84; 167pp; Japanese.

Polynucleotide sequences AAF74187 - AAF74218 encode proteins
AAB80608 - AAB80639, which impart environmental stress resistance
invention relates to a method for identifying DNA encoding proteins
cell, culturing the transformants in a salt-resistant organism into a host
untransformed host does not grow well, and selecting for viable clones.
The method is useful for obtaining DNA encoding environmental stress
resistance factors. The DNA encoding proteins conferring environmental
stress resistance, can be used in the production of plants resistant to
environmental stress, which can be cultivated in unfavourable
environments such as deserts, salt damaged ground, cold regions and the
oceans. They can be used for increasing the area of land covered by green

CC plants, and desert greening and afforestation, in order to counter the
CC effects of the increase in atmospheric carbon dioxide concentration. PCR
CC primers AAF74219 and AAF74220 are used in an example illustrating the
CC method of the invention.
XX Sequence 79 AA;
Query Match
Best Local Similarity 57.1%; Score 48; DB 22; Length 79;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 1 VILGVGPAKIOPEG 14
40 LVLGVPRAHPEG 53

RESULT 3
AAB31012
ID AAB31012 standard; Peptide; 28 AA.
AC AAB31012;
DT 01-FEB-2002 (first entry)
DE Peptide #3663 encoded by breast cell single exon nucleic acid probe.
KW Human; microarray; single exon probe; gene expression; breast;
disease; cancer.
OS *Homo sapiens*.
XX WO200157271-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00662.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0532366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,
breast, comprises number of single exon nucleic acid probes -
Claim 27; SEQ ID NO 13980; 327pp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon
nucleic acid probes for measuring gene expression in a sample derived
from human breast and BR 474 cells. The method involves contacting
the probes with a collection of detectably labelled nucleic acids
derived from mRNA of human breast, and then measuring the label
bound to each probe of the microarray. The probes are useful for
verifying the expression of regions of genomic DNA predicted to
code proteins. They are useful for gene discovery, and for
determining predisposition and/or prognostic breast disease. Gene
expression analysis is useful for assessing the toxicity of chemical
agents on cells. The microarray of this invention presents a far greater
diversity of probes for measuring gene expression, with far less bias
than expressed sequence tag microarrays. The method is suitable for
rapid production of functional information from genomic sequence. The
present sequence is a peptide encoded by a single exon nucleic acid
probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA;

Query Match 55.7%; Score 39; DB 22; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILVGPAKIOFE 13
: | | | | : | | |
DB 7 VLAVGPAOLQAE 18

RESULT 4
ABB36201
ID ABB36201 standard; Peptide; 28 AA.

AC ABB36201;

DE 04-FEB-2002 (first entry)

PE Peptide #3707 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-483447/52.

PS Claim 27; SEQ ID NO 28836; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA;

Query Match 55.7%; Score 39; DB 22; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILVGPAKIOFE 13
: | | | | : | | |
DB 7 VLAVGPAOLQAE 18

RESULT 5
ABB21577
ID ABB21577 standard; Protein; 28 AA.

AC ABB21577;

DE 23-JAN-2002 (first entry)

PE Protein #3576 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-488899/53.

PS Claim 15; SEQ ID NO 23347; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 28 AA;

Query Match 55.7%; Score 39; DB 22; Length 28;
Best Local Similarity 58.3%; Pred. No. 6.8;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILVGPAKIOFE 13
: | | | | : | | |
DB 7 VLAVGPAOLQAE 18

RESULT 6

AAAM56983
ID AAAM56983 standard; Protein; 28 AA.

AC AAAM56983;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29088.

KW Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI: 2001-483446/52.

DR Single exon nucleic acid probes for analyzing gene expression in human

XX brains -

XX Example 4; SEQ ID NO: 29088; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.

XX Sequence 28 AA;

XX Query Match

XX Best Local Similarity 55.7%; Score 39; DB 22; Length 28;

XX Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

XX 2 ILGVCPAKIOFE 13

XX : | | | | | : | |

XX 7 VLAVGPAQLQAE 18

XX

XX

PF 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI: 2001-488900/53.

DR Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 29675; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 28 AA;

XX Query Match

XX Best Local Similarity 55.7%; Score 39; DB 22; Length 28;

XX Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

XX 2 ILGVCPAKIOFE 13

XX : | | | | | : | |

XX 7 VLAVGPAQLQAE 18

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI: 2001-488901/53.

DR Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 29675; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 28 AA;

XX Query Match

XX Best Local Similarity 55.7%; Score 39; DB 22; Length 28;

XX Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

XX 2 ILGVCPAKIOFE 13

XX : | | | | | : | |

XX 7 VLAVGPAQLQAE 18

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

	from human placenta. The probes are useful for antenatal diagnosis of
CC	human genetic disorders.
XX	
Sequence	28 AA;
Query Match	55.7%; Score 39; DB 22; Length 28;
Best Local Similarity	58.3%; Pred. No. 6.8;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY	
2 ILGVPKIQIE 13	
: :	
Db	7 VLAAGPAQLQAE 18
RESULT 10	
AAM04898	
ID	AAM04898 standard; Protein; 28 AA.
XX	
AAM04898;	
AC	
DT	09-OCT-2001 (first entry)
XX	
Peptide #3580 encoded by probe for measuring breast gene expression.	
DE	
Probe; human; breast disease; breast cancer; development disorder;	
KW	Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX	
Homo sapiens.	
XX	
WC200157270-A2.	
PN	
09-ANG-2001.	
PD	
29-JAN-2001; 2001MO-US000661.	
PF	
XX	
04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-063366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
(MOL-) MOLECULAR DYNAMICS INC.	
PA	
Penn SG, Hanzel DK, Chen W, Rank DR;	
PI	
WP1: 2001-476286/51.	
DR	
Novel single exon nucleic acid probe used to measuring gene expression	
PT	in a human breast -
PS	
Claim 27; SEQ ID NO 13638; 322pp; English.	
XX	
The present invention relates to novel single exon nucleic acid probes	
XX	(see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC	such probe. The probes are useful for measuring human gene expression in
CC	a human breast sample, where the probe hybridises at high stringency to a
CC	nucleic acid expressed in the human breast. The probes are useful for
CC	predicting, diagnosing, grading, staging, monitoring and prognosing
CC	diseases of the human breast, particularly those diseases with polygenic
CC	aetiology. The diseases include: breast cancer, disorders of development,
CC	inflammatory diseases of the breast, fibrocystic changes, proliferative
CC	breast disease and non-carcinoma tumours.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
Sequence	28 AA;
Query Match	55.7%; Score 39; DB 22; Length 28;
Best Local Similarity	58.3%; Pred. No. 6.8;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ILGVGPAKIOFE 13
 :1 |||||:1
 Db 7 VLAVGPAQLOAE 18

RESULT 11

ABG38985
 ID ABG38985 standard; Peptide: 28 AA.

AC ABG38985;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 28650.

XX Human: single exon probe; asthma; lung cancer; COPD; ILD;
 XX chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

PD 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SC, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Claim 27; SEQ ID No 28650; 634pp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression to a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, Karagener syndrome,
 CC pulmonary alveolar proteinosis, lymphangioleiomyomatosis,
 CC pulmonary histiocytosis, primary ciliary dyskinesia, fibrocystic
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 28 AA:

Query Match 55.7%; Score 39; DB 23; Length 28;
 Best Local Similarity 58.3%; Pred. No. 6.8;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ILGVGPAKIOFE 13
 :1 |||||:1
 Db 7 VLAVGPAQLOAE 18

RESULT 12

AA28710
 ID AA28710 standard; Protein: 285 AA.

AC AA28710;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 34031.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161922.
 PR 29-OCT-1999; 99US-0162142.

Query Match 55.7%; Score 39; DB 21; Length 285;
 Best Local Similarity 42.9%; Pred. No. 81;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 OY 1 VIIGVPAKIOFEG 14
 Db 20 VVGGGPTGVFEG 33

RESULT 13
 AAC48146
 ID AAC48146 standard; Protein: 508 AA.
 AC AAC48146;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 60770.

KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 PN EPI033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX

PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-012180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132487.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 07-MAY-1999; 99US-0132486.
 PR 11-MAY-1999; 99US-0132487.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135333.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137528.
 PR 08-JUN-1999; 99US-0137724.
 PR 10-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139889.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140623.
 PR 29-JUN-1999; 99US-0140623.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142977.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144352.
 PR 21-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151348.
 PR 01-SEP-1999; 99US-0151830.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159325.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161992.
 PR 29-OCT-1999; 99US-0162142.

Query Match 55.7%; Score 39; DB 21; Length 508;
 Best Local Similarity 42.9%; Pred. No. 1.5e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 VILVGPAKIOPEG 14
 Db 243 VVVGSGPTGVFESG 256

RESULT 14
 ABB91938
 ID ABB91938 standard; Protein: 508 AA.

AC ABB91938;
 DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1149.
 DE Herbicidally active polypeptide SEQ ID NO 1149.

OS Arabidopsis thaliana.
 PN WO200210210-A2.
 PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.
 PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.
 PI Tietjen K, Weidler M;

DR WPI, 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

PS Claim 5; SEQ ID NO 1149; 261pp + Sequence listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

SQ Sequence 508 AA;

Query Match 55.7%; Score 39; DB 23; Length 508;
 Best Local Similarity 42.9%; Pred. No. 1.5e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 VILVGPAKIOPEG 14
 Db 243 VVVGSGPTGVFESG 256

```

RESULT 15
ABB90889
ID ABB90889 standard; Protein: 510 AA.
XX
AC ABB90889;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 100.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN MO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001MO-EP09892.
XX
PR 28-AUG-2001; 2001MO-EP09892.
XX
PA (PARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX
PS Claim 5; SEQ ID NO 100; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 510 AA;

```

Query Match 55.78; Score 39; DB 23; Length 510;
 Best Local Similarity 42.9%; Pred. No. 1.5e+02;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 VIUGVPAKIOPEG 14
   |::|||::||
Db 245 VVVGSGPTGVFSG 258

```

Search completed: January 13, 2003, 09:55:40
 Job time : 33.9487 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 12.9231 Seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVGPAKIOFEG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	75	2	T12188
2	70	100.0	75	2	S09098
3	64	91.4	75	2	S37240
4	56	80.0	77	2	T12187
5	53	75.7	74	1	SMBH1
6	53	75.7	77	2	S37239
7	49	70.0	79	2	J02128
8	43	61.4	78	2	T10087
9	42	60.0	78	2	S48038
10	40	57.1	82	2	P70569
11	40	57.1	82	2	T07076
12	40	57.1	82	2	T07114
13	40	57.1	212	2	F91002
14	40	57.1	215	2	E85847
15	40	57.1	456	2	AE0164
16	39	55.7	220	1	B5AG58
17	39	55.7	220	1	A13248
18	39	55.7	360	2	C84413
19	39	55.7	508	2	T02486
20	39	55.7	512	2	H86206
21	39	55.7	813	2	H86206
22	38	54.3	77	2	S52636
23	38	54.3	82	2	T03127
24	38	54.3	237	2	D71199
25	38	54.3	249	2	A48325
26	38	54.3	390	1	G69303
27	38	54.3	390	1	S75876
28	38	54.3	430	2	C97160
29	38	54.3	915	2	S44797

30	38	54.3	2632	2	T18718	dynein heavy chain
31	37	52.9	248	2	A70466	transcription anti
32	37	52.9	304	2	S13533	mRNA splice defec
33	37	52.9	314	2	S55179	mRNA splice defec
34	37	52.9	334	2	A36707	protein-glutamate
35	37	52.9	350	2	AD3560	methylated-DNA-lpr
36	37	52.9	371	2	T49100	hypothetical prote
37	37	52.9	375	1	S66272	alcohol dehydrogen
38	37	52.9	392	2	T51172	acetyl-CoA C-acety
39	37	52.9	406	2	T05865	hypothetical prote
40	37	52.9	421	2	A12824	NADH dehydrogenase
41	37	52.9	438	2	H97602	probable NADH dehy
42	37	52.9	447	2	T09414	pectinesterase hom
43	37	52.9	458	1	C70058	conserved hypothet
44	37	52.9	508	2	H90242	glycine dehydrogen
45	37	52.9	554	2	B85072	hypothetical prote

ALIGNMENTS

RESULT 1

T12188
metallothionein - fava bean
C:Species: *Vicia faba* (fava bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12188
R:Foley, R.C.; Liang, Z.M.; Singh, K.B.
Plant Mol. Biol. 33, 583-591, 1997
A:Title: Analysis of type 1 metallothionein cDNAs in *Vicia faba*.
A:Reference number: 217444; MUID:97238465; PMID:9132050
A:Accession: T12188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-75 <POL>
A:Cross-references: EMBL:X91078; NID:91150656; PIDN:CAA62552.1; PID:91150657
A:Experimental source: leave
A:Genetics: mlb
A:Gene: mtlb
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 100.0%; Score 70; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VILGVGPAKIOFEG 14
Db 36 VILGVGPAKIOFEG 49

RESULT 2

S09098
metallothionein - garden pea
C:Species: *Pisum sativum* (garden pea)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: S09098; S20347; S60719
R:Evans, I.M.; Gatehouse, L.N.; Gatehouse, J.A.; Robinson, N.J.; Croy, R.R.D.
FEBS Lett. 262, 29-32, 1990
A:Title: A gene from pea (*Pisum sativum* L.) with homology to metallothionein genes.
A:Reference number: S09098; MUID:90201367; PMID:2318309
A:Accession: S09098
A:Molecule type: DNA
A:Residues: 1-75 <EVA>
A:Cross-references: EMBL:Z23097; NID:9312504; PIDN:CAA80645.1; PID:9312505
R:Kille, P.; Winge, D.R.; Hatwood, J.L.; Kay, J.
FEBS Lett. 295, 171-175, 1991
A:Title: A plant metallothionein produced in *E. coli*.
A:Reference number: S20347; MUID:92111733; PMID:1765150
A:Accession: S20347
A:Molecule type: mRNA
A:Residues: 1-75 <KIT>
A:Cross-references: GB:Z23097; NID:9312504; PIDN:CAA80645.1; PID:9312505

A:Experimental source: cv. Feltham First
A:Accession: S60719
A:Molecule type: protein
A:Residues: 2-4;26-28;39-42;44-45;47-48;54-55 <K1W>
A:Experimental source: cv. Feltham First
C:Genetics: 17/2
C:Superfamily: metallochionein
C:Keywords: metal binding

Query Match
Best Local Similarity 100.0%; Score 70; DB 2; Length 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 VILGVPKIOPEG 14
DB 36 VILGVPKIOPEG 49

RESULT 3

S37240
metallochionein-like protein - white clover
C:Species: Trifolium repens (white clover)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
R:Ellison, N.W.
Submitted to the EMBL Data Library, September 1993.
A:Description: Sequence analysis of two cDNA clones for metallochionein-like proteins fr
A:Reference number: S37239
A:Accession: S37240
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-75 <ELL>
A:Cross-references: EMBL:226493; NID:9403328; PIDN:CAA81265.1; PID:9403329
C:Superfamily: metallochionein

Query Match
Best Local Similarity 100.0%; Score 64; DB 2; Length 75;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 VILGVPKIOPEG 13
DB 36 VILGVPKIOPEG 48

RESULT 4

T12187
metallochionein, type 1 - fava bean
C:Species: Vicia faba (fava bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12187
R:Polley, R.C.; Liang, Z.M.; Singh, K.B.
Plant Mol. Biol. 33, 583-591, 1997
A:Title: Analysis of type 1 metallochionein cDNAs in Vicia faba.
A:Reference number: Z17444; MUID:97238465; PMID:9132050
A:Accession: T12187
A:Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: mRNA
A:Residues: 1-77 <POL>
A:Cross-references: EMBL:X91077; NID:91150654; PIDN:CAA62551.1; PID:91150655
A:Experimental source: leaves
C:Superfamily: metallochionein
C:Keywords: metal binding

Query Match
Best Local Similarity 80.0%; Score 56; DB 2; Length 77;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Y 1 VILGVPKIOPEG 14
DB 39 VILGVPKIOPEG 52

RESULT 5

SMBH1
metallochionein - barley
C:Species: Hordeum vulgare (barley)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 22-Jun-1999
C:Accession: S17299; S15558
R:Okumura, N.; Nishizawa, N.K.; Umebara, Y.; Mori, S.
Plant Mol. Biol. 17, 531-533, 1991
A:Title: An iron deficiency-specific cDNA from barley roots having two homologous cys
A:Reference number: S17299; MUID:91355948; PMID:1832055
A:Accession: S17299
A:Molecule type: mRNA
A:Residues: 1-74 <OKU>
A:Cross-references: EMBL:X58540; NID:919006; PIDN:CAA41432.1; PID:919007
C:Genetics:
A:Gene: ids-1
C:Superfamily: metallochionein
C:Keywords: metal binding

Query Match
Best Local Similarity 75.7%; Score 53; DB 1; Length 74;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Y 1 VILGVPKIOPEG 13
DB 37 VILGVPKIOPEG 49

RESULT 6

S37239
metallochionein-like protein - white clover
C:Species: Trifolium repens (white clover)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
R:Ellison, N.W.
Submitted to the EMBL Data Library, September 1993
A:Description: Sequence analysis of two cDNA clones for metallochionein-like proteins
A:Reference number: S37239
A:Accession: S37239
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-77 <ELL>
A:Cross-references: EMBL:226492; NID:9403326; PIDN:CAA81264.1; PID:9403327
C:Superfamily: metallochionein

Query Match
Best Local Similarity 75.7%; Score 53; DB 2; Length 77;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Y 1 VILGVPKIOPEG 14
DB 40 VILGVPKIOPEG 53

RESULT 7

J02128
metallochionein - soybean
C:Species: Glycine max (soybean)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Mar-1995
C:Accession: J02128
R:Kawashima, I.; Inokuchi, Y.; Chino, M.; Kimura, M.; Shimizu, N.
Plant Cell Physiol. 32, 913-916, 1991
A:Title: Isolation of a gene for a metallochionein-like protein from soybean.
A:Reference number: J02128
A:Accession: J02128
A:Molecule type: mRNA
A:Residues: 1-79 <KAW>
A:Experimental source: seedling, cv. Saxa
C:Comment: This protein participates in detoxification and metabolism of heavy metals
C:Superfamily: metallochionein
C:Keywords: metal binding

Query Match
Best Local Similarity 70.0%; Score 49; DB 2; Length 79;

Best Local Similarity 57.1%; Pred. No. 0.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VILGVPKIOFEG 14
:|||||
DB 40 LVGVAPKAKOFE 53

RESULT 8

T10087
metallothionein - castor bean
C:Species: Ricinus communis (castor bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T10087
R:Yangala, S.; Bailey-Serres, J.
Plant Physiol. 109, 721, 1995
A:Title: Nucleotide sequence of a maize (Zea mays L.) cDNA (Accession No. U29383) coding
A:Reference number: Z16941; MUID:96030280; PMID:7480354
A:Accession: T10087
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-80 <VAN>
A:Cross-references: EMBL:L02306; NID:g169712; PID:g169713
A:Experimental source: strain Carmencita; tissue-type cotyledon
C:Genetics:
A:Gene: MTI
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 61.4%; Score 43; DB 2; Length 80;
Best Local Similarity 57.1%; Pred. No. 1.3;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VILGVPKIOFEG 14
:|||||
DB 40 LVGVAPKAKOFE 53

RESULT 9

S48038
metallothionein-like protein - kiwi fruit
C:Species: Actinidia chinensis var. deliciosa (kiwi fruit)
C:Date: 26-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999
C:Accession: S48038
R:Leeder, S.E.; Gardner, R.C.
Plant Mol. Biol. 25, 877-886, 1994
A:Title: Cloning and characterization of five cDNAs for genes differentially expressed
A:Reference number: S48035; MUID:94355660; PMID:8075403
A:Accession: S48038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-78 <LED>
A:Cross-references: EMBL:L27813; NID:9450244; PIDN:AAA53074.1; PID:9450245
C:Superfamily: metallothionein

Query Match 60.0%; Score 42; DB 2; Length 78;
Best Local Similarity 57.1%; Pred. No. 1.9;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VILGVPKIOFEG 14
:|||||
DB 40 LVGVAPKAKOFE 53

RESULT 10

F70569
hypothetical protein RV3489 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70569
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70569
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-54 <COL>
A:Cross-references: GB:Z95390; GB:AL123456; NID:g3261766; PIDN:CAB08712.1; PID:e31608
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3489

Query Match 57.1%; Score 40; DB 2; Length 54;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 ILGVGPARTO 11
:|||||
DB 44 MGIPTAKIE 53

RESULT 11

T07076
metallothionein type II B - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Apr-2000
C:Accession: T07076
R:Whitelaw, C.A.; Lehuquet, J.A.; Thurman, D.A.; Tomsett, A.B.
submitted to the EMBL Data Library, July 1996
A:Description: The isolation and characterization of type II metallothionein-like gene
A:Reference number: Z15900
A:Accession: T07076
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-82 <WHD>
A:Cross-references: EMBL:L77966; NID:g1449137; PIDN:AAB04675.1; PID:g1449138
A:Experimental source: strain Ailisa craig
C:Genetics:
A:Gene: MTB

A:Introns: 22/2
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 57.1%; Score 40; DB 2; Length 82;
Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VILGVPKIOF 12
:|||||
DB 41 LVGVGPKEKTSF 52

RESULT 12

T07114
metallothionein-like protein - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
C:Accession: T07114
R:Glitch, A.; Ganai, M.; Stephan, U.W.; Baumelein, H.
Plant Mol. Biol. 37, 701-714, 1998
A:Title: Structure, expression and chromosomal localization of the metallothionein-1
A:Reference number: Z15923; MUID:98349862; PMID:9687073
A:Accession: T07114
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-82 <GIR>

A:Cross-references: EMBL:Z68138; NID:g1103688; PIDN:CAA92243.1; PID:g1103689
A:Experimental source: cultivar Bonner Best; root
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 57.1%; Score 40; DB 2; Length 82;

Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LVIGVPAKIOF 12
Db 41 LVIGVPEKTSF 52

RESULT 13

F91002

Probable prophage repressor CI [imported] - Escherichia coli (strain O157:H7, substrain C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Jun-2002

C:Accession: F91002

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Khara, S.; Shiba, T.; Hattori, M.; Shihagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F91002

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-212 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA36413.1; PID:913362459; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECS2990

C:Superfamily: repressor protein ci

Query Match

Best Local Similarity 57.1%; Score 40; DB 2; Length 212;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVGPAPKIOF 12
Db 57 LGVSPAKIMF 66

RESULT 14

E85847

hypothetical protein 23358 [imported] - Escherichia coli (strain O157:H7, substrain EDL9 C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 03-Jun-2002

C:Accession: E85847

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85847

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <STO>

A:Cross-references: GB:AE005174; NID:912516411; PIDN:AGS7241.1; GSPDB:GN00145; UWGP:233
A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: 23358

C:Superfamily: repressor protein ci

Query Match

Best Local Similarity 57.1%; Score 40; DB 2; Length 215;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVGPAPKIOF 12
Db 60 LGVSPAKIMF 69

RESULT 15

AE0164

Probable exported protein YF01347 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AE0164
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilgall, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AE0164

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-456 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90176.1; PID:915979395; GSPDB:GN00175
C:Genetics:

A:Gene: YF01347

Query Match

Best Local Similarity 57.1%; Score 40; DB 2; Length 456;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LVIGVPAKIO 11
Db 8 LVIGAGPAPKIO 18

Search completed: January 13, 2003, 09:57:06
Job time: 14.9231 secs

Mon Jan 13 10:35:09 2003

us-09-554-941-1.rsp

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 ; Search time 14 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: us-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIOPEG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	70	100.0	MT1B_VICFA	Q41670 vicia faba
2	70	100.0	MT1A_PEA	P20830 pisum sativ
3	64	91.4	MTB_TRIP	P43399 trifolium r
4	60	85.7	MT1A_CICAR	Q3458 cicor ariet
5	56	80.0	MT1A_VICFA	Q41669 vicia faba
6	53	75.7	MT1A_HORVU	P26571 hordeum vul
7	53	75.7	MTA_TRIP	P43398 trifolium r
8	43	61.4	MT2_RICCO	P30564 ricinus com
9	42	60.0	MT2_RICCH	P43390 actinidia c
10	41	58.6	MT2_CICAR	Q39459 cicor ariet
11	41	58.6	A4M1_MOUSE	Q91K77 mus musculu
12	40	57.1	MT2B_LYCES	Q40158 lycopersico
13	40	57.1	VIB5_AGR5	P17795 agrobacteri
14	39	55.7	MT2_VICFA	Q41657 vicia faba
15	38	54.3	MT21_ORYSA	P94029 oryza sativ
16	38	54.3	6PGD_PIG	P14332 sus scrofa
17	38	54.3	SAT1_STYNY	P74241 synecocyst
18	38	54.3	YLS4_CABEL	P34389 caenorhabdi
19	38	54.3	NUSG_AOUAE	O67757 aquilex aeo
20	37	52.9	MRS4_YEAST	P23500 saccharomyc
21	37	52.9	MRS3_YEAST	P10566 saccharomyc
22	37	52.9	FRZG_MYXXA	P31758 myxococcus
23	37	52.9	ADH1_APTAU	P46645 apteryx aus
24	37	52.9	PME_MEDSA	Q42920 medicago sa
25	37	52.9	LRG1_YEAST	P35688 saccharomyc
26	37	52.9	P79331	P79331 b adams-2
27	37	52.9	ATG2_HUMAN	O95450 h adams-2
28	37	52.9	ATG2_HUMAN	O95450 h adams-2
29	36	51.4	MCA2_DROME	O94X99 mus musculu
30	36	51.4	TIN2_MOUSE	P05455 homo sapien
31	36	51.4	LA_HUMAN	P52209 homo sapien
32	36	51.4	6PGD_HUMAN	P26204 trifolium r
33	36	51.4	BGIS_TRIP	

34	36	51.4	803	1	Z151_HUMAN	O13105 homo sapien
35	35	50.0	78	1	MT2_MUSAC	O22319 musa acumin
36	35	50.0	344	1	COMC_METUA	O58820 methanococc
37	35	50.0	366	1	PHD1_YEAST	P36093 saccharomyc
38	35	50.0	374	1	PME_PERIN	O43043 petunia int
39	35	50.0	443	1	DCUA_HELPJ	O921C0 helicobacte
40	35	50.0	443	1	DCUA_HELPJ	O25425 helicobacte
41	35	50.0	458	1	OOAL_RHIME	O92Xp4 rhizobium m
42	35	50.0	463	1	STHA_PSEAE	P57112 pseudomonas
43	35	50.0	529	1	GUAA_MYCLE	P46810 mycobacteri
44	35	50.0	621	1	DCTB_RHIME	P13633 rhizobium m
45	35	50.0	643	1	NOSZ_ALCEU	O59105 alcaaligenes

ALIGNMENTS

RESULT 1	ID	MT1B_VICFA	STANDARD	PRT	75 AA
AC	Q41670				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Metallothionein-like protein 1B.				
GN	MT1B				
OS	Vicia faba (Broad bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;				
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.				
OX	NCBI_TaxID=3906;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=leaf;				
RX	MEDLINE=97238465; PubMed=9132050;				
RA	Foley R.C., Liang Z.M., Singh K.B.;				
RT	"Analysis of type 1 metallothionein cDNAs in Vicia faba.";				
RL	Plant Mol. Biol. 33:583-591(1997).;				
CC	-1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE				
CC	-1- RESIDUES THAT BIND VARIOUS HEAVY METALS.				
CC	-1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.				
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation its				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL; X91078; CAA62552.1;				
DR	InterPro; IPR000347; Metallothion_15.				
DR	Pfam; PF01439; Metallothio_2; 1.				
DR	ProDom; PD001611; Metallothion_15; 1.				
KW	Metal-binding; Metal-thiolate cluster; Multigene family.				
SQ	SEQUENCE 75 AA; 7750 MW; F59FC15A6A025BAA CRC64;				
Query Match	100.0%; Score 70; DB 1; Length 75;				
Best Local Similarity	100.0%; Pred. No. 1.2e-05;				
Matches 14; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
OY	1 VILGVPKIOPEG 14				
DB	36 VILGVPKIOPEG 49				
RESULT 2					
ID	MT1_PEA	STANDARD	PRT	75 AA	
AC	P20830				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				

DE Metallothionein-like protein 1.
 GN MTA.
 OS Pisum sativum (Garden pea).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 RN NCBI_TaxID=3888;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Feltham First; TISSUE=Root;
 RA MEDLINE=90201367; PubMed=2318309;
 RA Evans I.M., Gatehouse L.N., Gatehouse J.A., Robinson N.J.,
 RA Croy R.R.D.;
 RT "A gene from pea (Pisum sativum L.) with homology to metallothionein
 genes";
 RL FEBS Lett. 262:29-32(1990).
 RM [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9211733; PubMed=1765150;
 RA Kille P., Winge D.R., Harwood J.L., Kay J.;
 RT "A plant metallothionein produced in E. coli";
 RL FEBS Lett. 295:171-175(1991).
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z23097; CAAB0645.1; -
 DR PIR: S09098; S09098.
 DR PIR: S20347; S20347.
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 75 AA; 7608 MW; 68856F6F6883A70 CRC64;
 Query Match
 Best Local Similarity 100.0%; Score 70; DB 1; Length 75;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VILGVPKAKIOPEG 14
 DB 36 VILGVPKAKIOPEG 49
 RESULT 3
 MTB_TRIRP
 ID MTB_TRIRP STANDARD; PRT; 75 AA.
 AC P43399;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein B.
 GN MTA.
 OS Trifolium repens (Creeping white clover).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
 RN NCBI_TaxID=3899;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Huia; TISSUE=Stolon node;
 RA Ellison N.W., White D.W.R.;
 RT "Isolation of two cDNA clones encoding metallothionein-like proteins
 from Trifolium repens L.";
 RL (in) Plant Gene Register PGR96-068.

CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z26493; CAAB1265.1; -
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster; Multigene family.
 SQ SEQUENCE 75 AA; 7659 MW; 64ACAMCC234F54FE CRC64;
 Query Match
 Best Local Similarity 91.4%; Score 64; DB 1; Length 75;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VILGVPKAKIOFE 13
 DB 36 VILGVPKAKIOFE 48
 RESULT 4
 MT1_CICAR
 ID MT1_CICAR STANDARD; PRT; 75 AA.
 AC Q39436;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein 1 (MT-1).
 OS Cicer arietinum (Chickpea) (Garbanzo).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 RN NCBI_TaxID=3827;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Etioolated epicotyl;
 RA Dopico B., Labrador E., Ullan R.V., Munoz F.J.;
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC -!- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X95708; CAAB5008.1; -
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 75 AA; 7610 MW; 9806F0CB0B3F2BC CRC64;
 Query Match
 Best Local Similarity 85.7%; Score 60; DB 1; Length 75;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VILGVPKAKIOFEG 14
 DB 36 VILGVPKAKIOFEG 49

RESULT 5
MTA_VICFA STANDARD: PRT: 77 AA.
ID MTA_VICFA
AC 041669: 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-like protein 1A.
GN MTA.
OS *Vicia faba* (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3906;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=leaf; PubMed=9132050;
RX MEDLINE=97238465; PubMed=9132050;
RA Foley R.C., Liang Z.M., Singh K.B.;
RT "Analysis of type I metallothionein cDNAs in *Vicia faba*.";
RI Plant Mol. Biol. 33:583-591(1997).
CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X91077; CAA62551.1;
CC InterPro: IPR000347; Metallothion_15.
CC Pfam: PF01439; Metallothio_2; 1.
CC ProDom: PD001611; Metallothion_15; 1.
CC Metal-binding; Metal-thiolate cluster; Multigene family.
CC SEQUENCE 77 AA: 8022 MW; 40CED53369EB8727 CRC64;
SO
Query Match 80.0%; Score 56; DB 1; Length 77;
Best Local Similarity 71.4%; Pred. NO. 0.0036; 2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 2;
OY 1 VILGSPAKIOFEG 14
DB 39 VILGSPAKIHFDS 52

RESULT 6
MTL_HORVU STANDARD: PRT: 74 AA.
ID MTL_HORVU
AC P26571;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-like protein 1.
GN IDS-1.
OS *Hordeum vulgare* (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triliceae; Hordeum.
OX NCBI_TaxID=4513;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ehimadaka No.1; TISSUE=Root;
RX MEDLINE=91355948; PubMed=1832055;
RA Okumura N., Nishizawa N.-K., Umehara Y., Mori S.;
RT "An iron deficiency-specific cDNA from barley roots having two
RT homologous cysteine-rich MT domains.";
RI Plant Mol. Biol. 17:531-533(1991).
CC -1- FUNCTION: POSSIBLY RELEVANT TO MUGINEIC ACID-FAMILY (MAS)

PHYTOGIDROPHORE PRODUCTION IN FE-DEFICIENT BARLEY ROOTS.
MAY HAVE A FUNCTION AT THE REGULATORY REGION OF MAS SYNTHETIC
GENES OR FE(II)-MAS TRANSPORTER GENE BY CONJUGATING WITH FE(2+)
LIKE FOR PROTEIN.
-1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X56540; CAA1432.1;
CC PIR: S15558; SMH1.
CC InterPro: IPR000347; Metallothion_15.
CC Pfam: PF01439; Metallothio_2; 1.
CC ProDom: PD001611; Metallothion_15; 1.
CC Metal-binding; Metal-thiolate cluster.
CC SEQUENCE 74 AA: 7469 MW; F86C8CC3F65901C6 CRC64;
SO
Query Match 75.7%; Score 53; DB 1; Length 74;
Best Local Similarity 69.2%; Pred. NO. 0.012; 1; Indels 0; Gaps 0;
Matches 9; Conservative 3; Mismatches 1;
OY 1 VILGSPAKIOFEG 13
DB 37 VILGSPAKIOFEG 49

RESULT 7
MTA_TRIIP STANDARD: PRT: 77 AA.
ID MTA_TRIIP
AC P43398;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-like protein A.
GN MTA.
OS *Trifolium repens* (Creeping white clover).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolium.
OX NCBI_TaxID=3899;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Huia; TISSUE=stolon node;
RA Ellison N.W., White D.W.R.;
RT "Isolation of two cDNA clones encoding metallothionein-like proteins
RT from *Trifolium repens* L.";
RI (in) Plant Gene Register per96-068.
CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z26492; CAA81264.1;
CC InterPro: IPR000347; Metallothion_15.
CC Pfam: PF01439; Metallothio_2; 1.
CC ProDom: PD001611; Metallothion_15; 1.
CC Metal-binding; Metal-thiolate cluster; Multigene family.
CC SEQUENCE 77 AA: 5943 MW; 5F4C163BA644B8D1 CRC64;
SO
Query Match 75.7%; Score 53; DB 1; Length 77;
Best Local Similarity 64.3%; Pred. NO. 0.012;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 1 VILGVPKIOFEG 14
 Db 40 LVIGVGAERARHEG 53

RESULT 8

MT2_RICCO STANDARD: PRT; 80 AA.
 AC P30364;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE Metallothionein-like protein type 2.
 GN MT1.
 OS Ricinus communis (Castor bean).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Eumetazoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucotyledons; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Weig A.; Komor E.;
 RT "Isolation of a class II metallothionein cDNA from Ricinus communis
 L.";
 RL (1) Plant Gene Register PGR95-066.
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L02306; AAC37473.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 80 AA; 7953 MW; AA7304254491A3B7 CRC64;

Query Match 61.4%; Score 43; DB 1; Length 80;
 Best Local Similarity 57.1%; Pred. No. 0.76;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 VILGVPKIOFEG 14
 Db 40 LVIGVGAERARHEG 53

RESULT 9

MT2_ACTCH STANDARD: PRT; 78 AA.
 AC P43390;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Metallothionein-like protein type 2 (Kiwifruit)
 OS Actinidia chinensis (Kiwifruit)
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Eumetazoa; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; Ericales; Actinidiaceae; Actinidia.
 OX NCBI_TaxID=3625;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN-ev. Deliclosa Hayward; TISSUB=FRUIT;
 RX MEDLINE=94355660; PubMed=8075403;
 RA Ledger S.E.; Gardner R.C.;

RT "Cloning and characterization of five cDNAs for genes differentially
 RT expressed during fruit development of kiwifruit (Actinidia deliciosa
 RT var. deliciosa).";
 RL Plant Mol. Biol. 25:877-886(1994).

CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN YOUNG FRUIT WITH REDUCED
 CC EXPRESSION IN THE LATER STAGES OF FRUIT DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L27813; AAA53074.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster.
 SQ SEQUENCE 78 AA; 7828 MW; E2FE9BE0FFD501D CRC64;

Query Match 60.0%; Score 42; DB 1; Length 78;
 Best Local Similarity 57.1%; Pred. No. 1.1;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 VILGVPKIOFEG 14
 Db 40 LVIGVGAERARHEG 53

RESULT 10

MT2_CICAR STANDARD: PRT; 79 AA.
 AC Q39459;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE Metallothionein-like protein 2 (WF-2).
 DE Metallothionein-like protein 2 (WF-2).
 OS Cicer arietinum (Chickpea) (Garbanzo).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Eumetazoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucotyledons; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_TaxID=3827;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA TISSUE=Etliolated epicotyl;
 RA Dopico B.; Labrador E.; Gilan R.V.; Munoz F.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC -1- RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 15.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X55709; CA65009.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 DR KW Metal-binding; Metal-thiolate cluster; Multigene family.
 SQ SEQUENCE 79 AA; 7947 MW; 0DBE6C8DBF356071 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 79;
 Best Local Similarity 50.0%; Pred. No. 1.7;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIQE 14
 40 LVMGVASKTQEG 53

RESULT 11
 A4M1_MOUSE STANDARD: PRT; 449 AA.
 AC 09JCK7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adapter-related protein complex 4 mu 1 subunit (Mu subunit of Ap-4)
 DE (Ap-4 adapter complex mu subunit) (Mu-adaptin-related protein 2) (mu-
 AP2) (mu4).
 GN AP4M1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Werner H., Nave K., A.;
 RT "Mus musculus adaptor-related protein complex Ap-4 mu4 subunit.";
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SUBUNIT OF NOVEL TYPE OF CLATHRIN-OR NON-CLATHRIN-
 ASSOCIATED PROTEIN COAT INVOLVED IN TARGETING PROTEINS FROM THE
 TRANS-GOLGI NETWORK (TGN) TO THE ENDOSOMAL-LYSOSOMAL SYSTEM (BY
 SIMILARITY).
 CC -1- SUBUNIT: ADAPTER-LIKE COMPLEX 4 (AP-4) IS AN HETEROETRAMER
 COMPOSED OF TWO LARGE CHAINS (EPSILON/AP4E1 AND BETA/AP4B1), A
 MEDIUM CHAIN (MU/AP4M1) AND A SMALL CHAIN (SIGMA/AP4S1) (BY
 SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE TRANS-GOLGI NETWORK (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF242858; AAF63513.1;
 DR MGD: MGI:1337063; AP4M1.
 DR InterPro: IPR001392; Clathrn_med.
 DR Pfam: PF00928; Adap_comp_sub; 1.
 DR PRINTS: PR00314; CLATHRNADPT.
 DR PROSITE: PS00990; CLAT_ADAPTOR_M_1; FALSE_NEG.
 DR PROSITE: PS00991; CLAT_ADAPTOR_M_2; FALSE_NEG.
 DR Coated pits; Endocytosis.
 KW Coated pits; Endocytosis.
 SO SEQUENCE 449 AA; 49509 MW; 579AEELI8255D04 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 449;
 Best Local Similarity 63.6%; Pred. No. 9;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVGPKIQE 13
 396 LGVGPASLSE 406

RESULT 12
 A4M1_HUMAN STANDARD: PRT; 453 AA.
 ID A4M1_HUMAN
 AC 000189; 090HK9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adapter-related protein complex 4 mu 1 subunit (Mu subunit of Ap-4)
 DE (Ap-4 adapter complex mu subunit) (Mu-adaptin-related protein 2) (mu-
 AP2) (mu4).
 GN AP4M1 OR MDAAP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RC MEDLINE=97165966; Pubmed-9013859;
 RA Wang X., Kilmann M.W.;
 RT "Identification of a fourth adaptor-related protein, mu-ARP1 and
 mu-ARP2.";
 RL FEBS Lett. 402:57-61(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RC MEDLINE=99365166; Pubmed-10436028;
 RA Hirst J., Bright N.A., Rous B., Robinson M.S.;
 RT "Characterization of a fourth adaptor-related protein complex.";
 RL Mol. Biol. Cell 10:2787-2802(1999).
 RN [3]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RA TISSUE=Spleen;
 RC Pubmed-11139587;
 RA Aguilar R.C., Boehm M., Gorskova I., Crouch R.J., Tomita K.,
 Salto T., Ono H., Bonifacio J.S.;
 RT "Signal-binding specificity of the mu4 subunit of the adaptor protein
 complex, Ap-4.";
 RL J. Biol. Chem. 276:13145-13152(2001).
 CC -1- FUNCTION: SUBUNIT OF NOVEL TYPE OF CLATHRIN-OR NON-CLATHRIN-
 ASSOCIATED PROTEIN COAT INVOLVED IN TARGETING PROTEINS FROM THE
 TRANS-GOLGI NETWORK (TGN) TO THE ENDOSOMAL-LYSOSOMAL SYSTEM.
 CC -1- SUBUNIT: ADAPTER-LIKE COMPLEX 4 (AP-4) IS AN HETEROETRAMER
 COMPOSED OF TWO LARGE CHAINS (EPSILON/AP4E1 AND BETA/AP4B1), A
 MEDIUM CHAIN (MU/AP4M1) AND A SMALL CHAIN (SIGMA/AP4S1).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE TRANS-GOLGI NETWORK.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHLY EXPRESSED IN TESTIS AND
 LOWLY EXPRESSED IN BRAIN AND LUNG.
 CC -1- DOMAIN: INTERACTS SPECIFICALLY WITH TYROSINE-BASED SORTING
 SIGNALS.
 CC -1- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Y08387; CAA69667.1;
 DR EMBL: AF155158; AAD43328.1;
 DR EMBL: AF020796; AAD25869.1;
 DR Genew: HGNC:574; AP4M1.
 DR MIM: 602296;
 DR InterPro: IPR001392; Clathrn_med.
 DR Pfam: PF00928; Adap_comp_sub; 1.
 DR PROSITE: PS00990; CLAT_ADAPTOR_M_1; FALSE_NEG.
 DR PROSITE: PS00991; CLAT_ADAPTOR_M_2; FALSE_NEG.
 DR Coated pits; Endocytosis.
 KW Coated pits; Endocytosis.
 FT CONFLICT 338 R -> G (IN REF. 2).
 FT CONFLICT 345 R -> Q (IN REF. 2).
 FT CONFLICT 400 L -> M (IN REF. 2).
 FT CONFLICT 417 S -> C (IN REF. 2).
 FT CONFLICT 417 S -> C (IN REF. 2).
 SO SEQUENCE 453 AA; 50005 MW; 138B13A8C91D3444 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 453;
 Best Local Similarity 63.6%; Pred. No. 9.1;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 3 LGVGPARIQF 13
 DB 400 LGVGPASISFE 410

RESULT 13
 ID MT2B_LYCES STANDARD; PRT; 82 AA.
 AC 040158; 043514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein-like protein type 2 B.
 GN MTB.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Ailsa Craig;
 RA Whitehead C.A., Lehuquet J.A., Thurman D.A., Tomsett A.B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Bonnet Beste; TISSUE=Root;
 RA Girlich A., Herdik A., Balzer H., Stephan U., Baumlein H.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L77966; AAB04675.1;
 DR EMBL; Z68138; CAA92243.1;
 DR InterPro; IPR000347; Metallothion_15.
 DR Pfam; PF01439; Metallothio_2; 1.
 DR Prodom; PD001611; Metallothion_15; 1.
 DR Metal-binding; Metal-thiolate cluster; Multiligene family.
 FT CONFLICT 17 17 G->D (IN REF. 2).
 SO SEQUENCE 82 AA; 8253 MW; 6B298C4915CE895F CRC64;
 Query Match 57.1%; Score 40; DB 1; Length 82;
 Best Local Similarity 58.3%; Pred. No. 2.6;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 LVLGCPARIQF 12
 DB 41 LVLGVPKTSF 52

OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90318324; PubMed=2370849;
 RA Kuldau G.A., de Vos G., Owen J., McAffrey G., Zambryski P.;
 RT "The vir operon of Agrobacterium tumefaciens pTiC58 encodes 11 open
 RT reading frames";
 RL Mol. Gen. Genet. 221:256-266(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90301800; PubMed=2194232;
 RA Rogowsky P.M., Powell B.S., Shirsu K., Lin T.-S., Morel P.,
 RA Zyprian E.M., Steck T.R., Kado C.I.;
 RT "Molecular characterization of the vir regulon of Agrobacterium
 RT tumefaciens: complete nucleotide sequence and gene organization of
 RL the 28.63-kbp regulon cloned as a single unit";
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Westler E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58";
 RL Science 294:2317-2323(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Seay C., Strub G.,
 RA Cleo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 CC -1- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL
 CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA
 CC TRANSFER TO PLANT CELLS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X53264; CAA37358.1;
 DR EMBL; J03320; AAB1555.1;
 DR EMBL; J03320; AAB1555.1;
 DR EMBL; AE009435; AAL46407.1;
 DR EMBL; AE009435; AAK90933.1;
 DR PIR; S12345; B5AG58.
 DR Crown gall tumor; Plasmid; Signal; Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 220 VIRB5 PROTEIN.
 SO SEQUENCE 220 AA; 23269 MW; 78FEFF913E03EA207 CRC64;
 Query Match 55.7%; Score 39; DB 1; Length 220;
 Best Local Similarity 63.6%; Pred. No. 10;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 2 LVLGCPARIQF 12

Db 15 LLSYGPAAQF 25

11111111

```

RESULT 15
MT2_VICFA STANDARD; PRT; 77 AA.
ID MT2_VICFA
AC 041657;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-like protein type 2.
GN MT1.
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OC NCBI_TaxID=3906;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95036014; PubMed=7948889;
RX Foley R.C., Singh K.B.;
RT "Isolation of a Vicia faba metallothionein-like gene: expression in
RT foliar trichomes."
RT Plant Mol. Biol. 26:435-444(1994).
CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE LEFT, STEM AND FLOWER, AT
CC VERY LOW LEVELS IN ROOTS AND IS NOT DETECTABLE IN MESOPHYLL
CC PROTOPLASTS.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 15.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X77254; CAA54471.1;
CC InterPro: IPR000347; Metallothion_15.
CC Pfam: PF01439; Metallothio_2; 1.
CC ProDom: PD001611; Metallothion_15; 1.
CC Metal-binding; Metal-thiolate cluster; Multigene family.
CC KW SEQUENCE 77 AA; 7730 MW; 7D7BAD8D023BF3C CRC64;
CC SO
Query Match 54.38; Score 38; DB 1; Length 77;
Best Local Similarity 53.88; Pred. No. 5.6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VILGVPKIQFE 13
:|:|:|:|:|:|
DB 40 LIMGVSEKAYE 52

```

Search completed: January 13, 2003, 09:54:05
Job time : 16.3333 secs

OM protein - protein search, using sw model
January 13, 2003, 09:53:15 : Search time 31.9487 seconds
(without alignments)
58.391 Million cell updates/second

```

Searched: 908470 segs, 100%
Total number of hits satisfying chosen parameters: 0
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Minimum Match 0%
Maximum Match 100%
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1964.DAT.*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1965.DAT.*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1966.DAT.*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1967.DAT.*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1968.DAT.*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1969.DAT.*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1970.DAT.*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1971.DAT.*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1972.DAT.*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1973.DAT.*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1974.DAT.*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1975.DAT.*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1976.DAT.*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1977.DAT.*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1978.DAT.*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1979.DAT.*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysts of the total score distribution.

SUMMARIES

[illegible]

4 1980-1981; Conservative; Pred. No. 7.4; Length 321;

CONFIDENTIAL 84

ABP03100 standard; Proct:

24-JUN-2002 (first entry)

Human OREX protein sequence

Human; open reading frame
hyperproliferative

cardiovascular disease; cancer; degenerative disorder; gene therapy; hyperproliferative disorder; ORFX; psoriasis; rheumatoid arthritis

hypertension; diabetes mellitus; haemorrhage;
osteoarthritis; neurodegenerative
disease; benign tumour; cirrhosis;

autoimmune disorder; immune deficiency; cholesterol ester storage; myasthenia; thyroidism; systemic lupus erythematosus; reactive disorder;

Homio sapiens
griavis.
tneumatoid arthritis; autoimmune disease;
infectious disease;
arthritis; autoimmunity

...ne lymphoiditis;

06-DEC-2002

29-MAY-2001.

2001; 2001WO-US10836
30-MAY-2000

29-AUG-2000; 2000US-206132P

(CURA-) CURAGEN CORP
228/16P.

Shimkets RA, Leach MD

WPI; 2002-106308/14
N-PSDB; 2002-106308/14

ABN18852.

... human polyp
... preventing and tre
... ybern...

proliferative

SECURITY; SEQ ID

present invention referred to as one of the

specific proteins given in a

... or prevent
... in humans

ences can be associated

... of cancer,

osteoarthritis, neuro
transplantation

erythematous, age disease

— 200, VAF 1

1

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. OPRX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 77 AA;

Query Match 58.9%; Score 43; DB 23; Length 77;
 Best Local Similarity 61.5%; Pred. No. 3.4;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 QASIEGRYAPQP 13
 : : : : :
 Db 53 KAWVEGRPAQP 65

RESULT 4

ID ABB57844 standard; Protein; 283 AA.

XX ABB57844;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 324.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL01947.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS

PS Disclosure; SEQ ID NO 324; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 283 AA;

Query Match 58.9%; Score 43; DB 22; Length 283;
 Best Local Similarity 63.6%; Pred. No. 15;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYAPQP 14
 : : : : :
 Db 93 IKGRIHAPQP 103

RESULT 5

ID ABB12235 standard; peptide; 105 AA.

XX ABB12235;

DT 11-JAN-2002 (first entry)

DE Human secreted protein homologue, SEQ ID NO:2605.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; vasotropic; cardiac; antiviral; antibacterial;
 KW antifungal; vulnary; antitumor.

XX Homo sapiens.

PN WO200157188-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US03800.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457740/49.

DR N-PSDB; ABA09479.

PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -

PS Claim 20; Page 318; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156456.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.5%; Score 42; DB 21; Length 313;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 QASIEGRYAPOP 13
Db 229 QEGAGHTEPOP 241

RESULT 7
AAG3807
ID AAG3807 standard; Protein; 474 AA.
XX
AC AAG3807;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27250.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131448.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136362.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140351.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 02-JUL-1999; 99US-0141842.
PR 06-JUL-1999; 99US-0142154.
PR 08-JUL-1999; 99US-0142055.
PR 09-JUL-1999; 99US-0142390.
PR 12-JUL-1999; 99US-0142803.
PR 13-JUL-1999; 99US-0142920.
PR 14-JUL-1999; 99US-0142977.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 19-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144633.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 05-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147305.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147938.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0151330.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154179.
PR 23-SEP-1999; 99US-0155139.
PR 28-SEP-1999; 99US-0155486.
PR 29-SEP-1999; 99US-0155659.
PR 04-OCT-1999; 99US-0156458.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.

Mon Jan 13 10:35:10 2003

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
Query Match 57.8%; Score 42; DB 21; Length 474;
Best Local Similarity 53.8%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 QASIEGRYTAPOP 13
DB 390 QEGLAGRTEPOP 402
RESULT 8
AG23806
ID AAG23806 standard; Protein: 493 AA.
AC AAG23806;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27249.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27249.
XX protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145276.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.

```

Query Match      56.8%  Score 41.5;  DB 22;  Length 69.
Best Local Similarity
Matches 8.2%  Pred no f c

```

```

Query Match          56.8%; Score 41.5; DB 22; Length 69;
Best Local Similarity 53.3%; Pred. No. 5.6;
Matches      8; Conservative 4; Mismatches 2; Indels 1; Gaps 1

```

OY 1 QASIE-GRYTAPOPQ 14
 ||:|:|:|:|
 Db 39 QAGVOMGSYSSPPQ 53

RESULT 10
 ABB39729
 ID ABB39729 standard; Peptide; 49 AA.
 AC ABB39729;

XX 04-FEB-2002 (first entry)

DT Peptide #7235 encoded by human foetal liver single exon probe.

DE Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 32364; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human foetal

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX foetal liver. The present sequence is a peptide encoded by a single exon

XX nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 49 AA;

OY 4 IEGRYTAPOP 13
 :||:|:|:|
 Db 20 VEGRHTVPRP 29

RESULT 11
 AAM73093
 ID AAM73093 standard; Protein; 49 AA.

XX AAM73093;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33399.
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 33399; 658pp + Sequence listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

XX Sequence 49 AA;

OY 4 IEGRYTAPOP 13
 :||:|:|:|
 Db 20 VEGRHTVPRP 29

RESULT 12
 AAM33312
 ID AAM33312 standard; Protein; 49 AA.

XX AAM33312;

XX 17-OCT-2001 (first entry)

XX Peptide #7349 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-48897/53.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 33581; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see A131315-A157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.

SQ Sequence 49 AA;

Query Match 56.2%; Score 41; DB 22; Length 49;
 Best Local Similarity 60.0%; Pred. No. 4.7;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 13
 DB 20 VEGRTVPRP 29

RESULT 13

ID ABG42939 standard; Peptide; 49 AA.

AC ABG42939;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 32604.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagamer syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW hyaline membrane disease.

OS Homo sapiens.

PM WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US00665.

PR 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

PS Claim 27; SEQ ID No 32604; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC sample derived from human lung; measuring gene expression in a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagamer syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 49 AA;

Query Match 56.2%; Score 41; DB 23; Length 49;
 Best Local Similarity 60.0%; Pred. No. 4.7;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 IEGRYTAPOP 13
 DB 20 VEGRTVPRP 29

RESULT 14

ID ABG07274 standard; Protein; 186 AA.

AC ABG07274;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #7265.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI, 2001-639362/73.
 XX DR N-PSDB; AAS71461.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID No 37633; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 186 AA;
 OY 3 SIEGRTAPQP 14
 DB 166 NVNGRPTAPQPK 177
 Query Match 56.2%; Score 41; DB 22; Length 186;
 Best Local Similarity 58.3%; Pred. No. 21;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 RESULT 15
 AAY22919
 ID AAY22919 standard; peptide; 11 AA.
 XX AC AAY22919;
 XX DT 19-AUG-1999 (first entry)
 XX DE Galactose oxidase linker used to make proteinase inhibitor fusions.
 XX KW Peptide linker; fusion protein; pathogen resistance; pathogen tolerance;
 XX KW plant; transgene; proteinase inhibitor; Arabidopsis thaliana;
 XX KW parasitic nematode.

OS Synthetic.
 XX PN WO9928484-A1.
 XX PD 10-JUN-1999.
 XX PF 01-DEC-1998; 98WO-EP07792.
 XX PR 03-DEC-1997; 97GB-0025556.
 XX PA (NOVS) NOVARTIS AG.
 XX PI Atkinson HJ, McPherson MJ, Urwin PE;
 XX DR WPI, 1999-385387/32.
 XX PT Proteinase inhibitor fusion proteins
 XX PS Claim 9; Page 23; 39pp; English.
 XX CC The present sequence represents a peptide linker used in the fusion
 CC proteins of the invention. The specification describes a method for
 CC improving pathogen resistance or tolerance of a plant. The method
 CC comprises transformation of the plant with a transgene encoding a
 CC fusion protein which comprises two or more protein or domains that
 CC are capable of improving pathogen resistance or tolerance when
 CC expressed on their own. Specifically, two distinct proteinase
 CC inhibitors are co-delivered, as a fusion, to Arabidopsis thaliana. The
 CC proteins or domains are connected by a peptide linker. The method is
 CC used to improve pathogen resistance or tolerance of a plant and its
 CC descendants, especially against parasitic nematode attack.
 XX SQ Sequence 11 AA;
 OY 1 QASIEGRTAPQP 14
 DB 1 QAS---STTAPQP 11
 Query Match 55.5%; Score 40.5; DB 20; Length 11;
 Best Local Similarity 71.4%; Pred. No. 1.1;
 Matches 10; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 Search completed: January 13, 2003, 09:55:42
 Job time : 32.9487 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 53.1282 Seconds

(without alignments)
54.296 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIQIEG 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriopl:*
17: sp_archaeopl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	92.9	75	10	Q9SP23 medicago sa
2	53	75.7	74	10	Q23812 hordeum vul
3	46	65.7	75	10	Q9LUX2 pyrus pyrif
4	43	61.4	77	10	Q93X22 quercus sub
5	43	61.4	79	10	Q9AUI6 typha latif
6	43	61.4	79	10	Q944W1 typha latif
7	42	60.0	81	10	Q941B7 atropa bell
8	42	60.0	602	10	Q9SC90 medicago tr
9	41	58.6	69	10	Q82562 elaeagnus u
10	41	58.6	80	10	Q9FR40 petunia hyb
11	41	58.6	453	4	Q8WV65 homo sapien
12	40	57.1	54	16	Q06352 mycobacteri
13	40	57.1	79	10	Q80335 citrus unsh
14	40	57.1	82	10	Q94198 atropa bell
15	40	57.1	201	16	Q92R37 rhizobium m
16	40	57.1	212	2	Q9EYB2 escherichia

17	40	57.1	215	16	Q8XSE1 escherichia
18	40	57.1	456	16	Q8ZGF4 versinia pe
19	39	55.7	76	10	Q9M4N0 persia amer
20	39	55.7	360	17	Q9HMA0 halobacteri
21	39	55.7	441	13	Q12969 gallus gall
22	39	55.7	495	10	Q9ST63 solanum tub
23	39	55.7	499	10	Q941V1 oryza sativ
24	39	55.7	508	10	Q80874 arabidopsis
25	39	55.7	512	10	Q9LML0 arabidopsis
26	39	55.7	813	16	Q9KGS2 bacillus ha
27	38	54.3	66	10	Q9ZSC6 ipomoea bat
28	38	54.3	115	17	Q8TKX4 methanosarc
29	38	54.3	167	16	Q98K40 rhizobium l
30	38	54.3	237	17	Q59554 pyrococcus
31	38	54.3	334	4	Q29818 archaeoglob
32	38	54.3	343	4	Q8TE01 actinobact
33	38	54.3	368	2	Q9LAA3 clostridium
34	38	54.3	396	16	Q8XN53 bradyrhizob
35	38	54.3	430	16	Q97HA1 clostridium
36	38	54.3	556	2	Q45232 infectious
37	38	54.3	616	12	Q98VU1 infectious
38	38	54.3	616	12	Q91BR3 infectious
39	38	54.3	616	12	Q91BR3 infectious
40	38	54.3	2632	5	P90736 P90736
41	38	52.9	260	16	Q987C0 rhizobium l
42	37	52.9	350	16	Q8YCX3 brucea me
43	37	52.9	371	10	Q9SUY4 arabidopsis
44	37	52.9	382	2	Q9ZG19 alcaigenes
45	37	52.9	392	2	Q9ZB53 alcaigenes

ALIGNMENTS

RESULT 1	ID	Q9SP23	PRELIMINARY:	PRT:	75 AA.
AC	Q9SP23	01-MAY-2000 (TREMUREL.13, Created)			
DT	01-MAY-2000 (TREMUREL.13, Last sequence update)				
DT	01-DEC-2001 (TREMUREL.19, Last annotation update)				
DE	Type 1 metallothionein.				
GN	MTL.				
OS	Medicago sativa (Alfalfa).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid 1; Fabiales; Fabaceae; Papilionoideae; Trifolieae; Medicago.				
OX	NCBI_TaxID=3879;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. LAHONTAN; TISSUE=ROOT;				
RA	PubMed=11448760;				
RX	Potenza C., Thomas S.H., Sengupta-Gopalan C.;				
RT	"Genes induced during early response to Meloidogyne incognita in roots				
RT	of resistant and susceptible alfalfa cultivars.";				
RL	Plant Sci. 161:289-299(2001).				
DR	EMBL: AF189766; AAF04584.1;				
DR	InterPro: IPR000347; Metallothion_15.				
DR	Plan: PF01439; Metallothion_15; 1.				
DR	Prodom: PD001611; Metallothion_15; 1.				
SO	SEQUENCE 75 AA; 7537 MW; 7872EF0F02BD22FB CRC64;				

Query Match
Best Local Similarity 92.9%; Score 65; DB 10; Length 75;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 VILGVPKIQIEG 14
36 VILGVPKIQIEG 49

RESULT 2
Q23812

ID 023812 PRELIMINARY; PRT; 74 AA.
 AC 023812;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Plant metallothionein-like protein.
 GN ID1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VAR. NK1558;
 RA Nakanishi H., Okumura N., Kanegae R., Umehara Y., Nishizawa N.,
 Mori S.;
 RT "A plant metallothionein-like gene from iron deficiency barley
 roots";
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D50641; BAA23628.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SO SEQUENCE 74 AA; 7409 MW; F86C8CC3ED58B1C6 CRC64;

Query Match 75.7%; Score 53; DB 10; Length 74;
 Best Local Similarity 69.2%; Pred. No. 0.054;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 13
 Db 37 IIVGVAPKMFEG 49
 :||||| |||

RESULT 3
 O9LUX2 PRELIMINARY; PRT; 75 AA.
 ID 09LUX2;
 AC 09LUX2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Metallothionein-like protein.
 GN PPRF016.
 OS Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX euroids 1; Rosales; Rosaceae; Maloideae; Pyrus.
 OX NCBI_TaxID=3767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. KIKUSUI; TISSUE=FRUIT;
 RA Itai A., Tanaka T., Tanabe K., Tamura F.;
 RT "Pyrus pyrifolia ripening associated mRNA, clone PPRF016";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021790; BAA96449.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SO SEQUENCE 75 AA; 7831 MW; A772B9B9367755A CRC64;

Query Match 65.7%; Score 46; DB 10; Length 75;
 Best Local Similarity 57.1%; Pred. No. 0.099;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 Db 36 IIVGVAPKMFEG 49
 :||| |||

RESULT 4
 O93X22 PRELIMINARY; PRT; 77 AA.
 ID 093X22

AC 093X22;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Metallothionein-like protein.
 GN MT.
 OS Quercus suber (Cork oak).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX euroids 1; Fagales; Fagaceae; Quercus.
 OX NCBI_TaxID=58331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PERIDERM;
 RA Mir G.;
 RT "Cork metallothionein";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ277599; CAC39481.2;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SO SEQUENCE 77 AA; 7673 MW; C867A295A9DAB0A5 CRC64;

Query Match 61.4%; Score 43; DB 10; Length 77;
 Best Local Similarity 57.1%; Pred. No. 3.5;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 Db 39 IIVGVAPKMFEG 52
 :||| |||

RESULT 5
 O9AUI6 PRELIMINARY; PRT; 79 AA.
 ID 09AUI6;
 AC 09AUI6;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Metallothionein-like protein.
 OS Typha latifolia (Common cattail).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
 OX NCBI_TaxID=4733;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.W., Ma M., Tam F.Y., Wong Y.S.;
 RT "Cloning and sequencing of metallothionein (MT) cDNA from a heavy
 metal tolerant plant, Typha latifolia";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF279655; AAK28022.1;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 SO SEQUENCE 79 AA; 7954 MW; BBI1C53D3E4EC2D CRC64;

Query Match 61.4%; Score 43; DB 10; Length 79;
 Best Local Similarity 64.3%; Pred. No. 3.6;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 VILGVPKIOFEG 14
 Db 40 IIVGVAPKMFEG 53
 :||| |||

RESULT 6
 O94AW1 PRELIMINARY; PRT; 79 AA.
 ID 094AW1;
 AC 094AW1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Type 2 metallothionein-like protein.

Mon Jan 13 10:35:10 2003

us-09-554-941-1.rpt

Page 3

OS Typha latifolia (Common cattail).
OC Spermatophyta: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; Liliopsida; Poales; Typhaceae; Typha.
OC NCBI_TaxID=4733;
[1]
RN SEQUENCE FROM N.A.
RP Zhang Y.-W., Tam F.-Y., Wong Y.-S.;
RA "Cloning and characterization of type 2 metallothionein-like gene from
RT Typha latifolia."
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF419840; AL09705.1; T15.
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 79 AA; 8053 MW; B8090030C6F1EC2D CRC64;
SQ
Query Match 61.4%; Score 43; DB 10; Length 79;
Best Local Similarity 64.3%; Pred. No. 3.6; 4; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 0;
DB 1 VILGVPKIQFEG 14
40 MVLGVAPEKGFEG 53
RESULT 7
Q94187 PRELIMINARY; PRT; 81 AA.
ID Q94187
AC Q94187
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Putative metallothionein-like protein type 2B.
DE Atriplex belladonna (belladonna) (Deadly nightshade).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Atriplex.
OC NCBI_TaxID=33113;
OX [1]
RN SEQUENCE FROM N.A.
RP Nouar E., Baucher M., Jaziri M.;
RA "Differential gene expression in Atriplex belladonna leafy gall."
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP Nouar E.;
RA Thesis (2001), Department of Plant Biotechnology,
RL Universite Libre de Bruxelles, Brussels, Belgium.
RL EMBL: AJ309387; CAC40757.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 81 AA; 8294 MW; 1CE236120CDBB9F4 CRC64;
SQ
Query Match 60.0%; Score 42; DB 10; Length 81;
Best Local Similarity 53.8%; Pred. No. 5.6; 3; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 0;
DB 1 VILGVPKIQFEG 13
40 LVLSVGPKEKTSFD 52
RESULT 8
Q9SC90 PRELIMINARY; PRT; 602 AA.
ID Q9SC90
AC Q9SC90
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE Pectin methyl-esterase PER precursor (PC 3.1.1.11).
DE Pectin methyl-esterase PER precursor (PC 3.1.1.11).
GN PER.
OS Medicago truncatula (Barrel medic).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
OC NCBI_TaxID=3880;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. JEMALONG;
RA Perez-Hormaeche J., Kondorosi A., Palomares A.J., Ratet P.;
RT "MPEP, a new pectin methyl-esterase gene from Medicago truncatula
RT involved in the infection process of root nodules by the symbiotic
RT host Sinorhizobium meliloti."
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. JEMALONG;
RC Perez-Hormaeche J., Rodriguez-Llorente I.D., Kondorosi A.,
RA Palomares A.J., Ratet P.;
RT "A genomic cluster of pectin methyl-esterase genes strongly suggest a
RT common evolutionary origin for isoforms induced in pollen and root
RT nodules development."
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ249611; CAB5290.2;
DR InterPro: IPR000070; Pectinesterase.
DR Pfam: PF01095; Pectinesterase; 1.
DR ProDom: PD00800; PECTINESTERASE_1; 1.
DR PROSITE: PS00012; PHOSPHOTANTERETINE; UNKNOWN_1.
DR SIGNAL 1
KW SIGNAL
FT SIGNAL
SQ SEQUENCE 602 AA; 65003 MW; FBE27E825FE62B2F CRC64;
POTENTIAL.
ID 082562
AC 082562
DT 01-NOV-1998 (TREMREL. 08, Created)
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE Metallothionein homolog.
DE Eleagnus umbellata (Autumn olive).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Elaeagnaceae; Eleagnus.
OC NCBI_TaxID=43233;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE-ROOT NODULE;
RA Kim H.B., An C.S.;
RT "Isolation and characterization of cDNA clone encoding metallothionein
RT from the root nodule of Eleagnus umbellata."
RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF091514; AAC62105.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothion_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
DR SEQUENCE 69 AA; 7141 MW; 2B98518FEB08938A CRC64;
SQ
Query Match 58.6%; Score 41; DB 10; Length 69;
Best Local Similarity 61.5%; Pred. No. 7.2; 4; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;
DB 2 VILGVPKIQFEG 14
31 VILGVAPEKGFEG 43

RESULT 10

09FR40
ID 09FR40 PRELIMINARY; PRT; 80 AA.
AC 09FR40: (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Putative metallothionein-like protein.
MT2.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OX NCBI_TaxID=4102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PETUNIA MULTIFLORA HYBRID, AND CV. PRIMETIME ROSE,
RA Lai C.P., Shaw J.F.;
RT Cloning and characterization of cDNAs for metallothionein-like
RL Submitted from Petunia.
DR EMBL: AF201384; AAG39645.1;
DR InterPro: IPR000347; Metallothion_15.
DR Pfam: PF01439; Metallothio_2; 1.
DR ProDom: PD001611; Metallothion_15; 1.
SQ SEQUENCE 80 AA; 8084 MW; 4248ADEOC05FACC7 CRC64;

Query Match

Best Local Similarity 58.6%; Score 41; DB 10; Length 80;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILGVSPAKIOF 12
Db 40 LILGVSPAKIOF 51

RESULT 11

08WY65
ID 08WY65 PRELIMINARY; PRT; 453 AA.
AC 08WY65:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-JUN-2002 (TREMBlrel. 20, Last sequence update)
OS Hypothetical 50.0 kDa protein.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC018705; AAH18705.1;
DR InterPro: IPR001392; Clathrm_med.
DR Pfam: PF00928; Adap_comp_sub; 1.
KW Hypothetical protein.
SQ SEQUENCE 453 AA; 49977 MW; AEDCA8GCAED08B7 CRC64;

Query Match

Best Local Similarity 58.6%; Score 41; DB 4; Length 453;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 LGVSPAKIOF 13
Db 400 LGVSPAKIOF 410

RESULT 12

006352
ID 006352 PRELIMINARY; PRT; 54 AA.

RESULT 13

080335
ID 080335 PRELIMINARY; PRT; 79 AA.
AC 080335:
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-MAR-2002 (TREMBlrel. 08, Last sequence update)
DE Metallothionein-like protein.
GN Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FRUIT;
RA MEDLINE=98267202; PubMed=9602134;
RA Moriyuchi T., Kita M., Hisada S., Endo-Inagaki T., Omura M.;
RT "Characterization of gene repertoires at mature stage of citrus fruits
through random sequencing and analysis of redundant metallothionein-
like genes expressed during fruit development.";

Query Match

Best Local Similarity 57.1%; Score 40; DB 16; Length 54;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 ILGVSPAKIO 11
Db 44 MGVSPAKIO 53

RL Gene 211:221-227(1998).
 DR EMBL: AB008100; BAA31561.1; -;
 DR InterPro: IPR000347; Metallothion_15.
 DR Pfam: PF01439; Metallothio_2; 1.
 DR ProDom: PD001611; Metallothion_15; 1.
 DR SEQUENCE 79 AA; 7741 MW; CDDAF5CC41BD36AE CRC64;

Query Match 57.1%; Score 40; DB 10; Length 79;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LVIGVPAKIQFEG 14
 :||| | :|||
 Db 42 LVIGVAPVKMHSEG 55

RESULT 14

Q94198 PRELIMINARY; PRT; 82 AA.

AC Q94198; 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)

DE Methallothioneine-like protein.

GN MT2B.

OS Atropa belladonna (Belladonna) (Deadly nightshade).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.

NCBI_TaxID=33113;

RP [1]

RP SEQUENCE FROM N.A.

RA Nouar E., Marie B., Jaziri M.;

RT "Differential gene expression in *Atropa belladonna* leafy gall."

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ297968; CAC40742.1; -;

DR InterPro: IPR000347; Metallothion_15.

DR Pfam: PF01439; Metallothio_2; 1.

DR ProDom: PD001611; Metallothion_15; 1.

DR SEQUENCE 82 AA; 8437 MW; 73A012103E4E9DF0 CRC64;

QY 1 LVIGVPAKIQF 12

Db 41 LVIGVPEKTSF 52

RESULT 15

Q92R37 PRELIMINARY; PRT; 201 AA.

AC Q92R37; 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Hypothetical protein R01087.

GN R01087 OR SMC02636.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

NCBI_TaxID=382;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE-21396507; PubMed-11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,

RA Renard G., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591786; CAC45666.1; -;
 KW Hypothetical protein, complete proteome.
 DR SEQUENCE 201 AA; 21985 MW; 6610A10C3F5DDC84 CRC64;

Query Match 57.1%; Score 40; DB 16; Length 201;
 Best Local Similarity 58.3%; Pred. No. 33;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LGVGPAPKIQFEG 14
 :||| | :|||
 Db 63 LGVGPAPKIQFEG 74

Search completed: January 13, 2003, 09:59:42
 Job time : 56.1282 secs

Mon Jan 13 10:35:08 2003

us-09-554-941-1.ra1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 : Search time 12.2051 Seconds
(without alignments)
33.750 Million cell updates/sec

Title: US-09-554-941-1
Perfect score: 70
Sequence: 1 VILGVPKIOFEG 14

Scoring table: BLSDM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/aa/PCITUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	70	100.0	75	1 US-08-322-962-5	Sequence 5, Appl
2	70	100.0	75	3 US-08-450-653-5	Sequence 5, Appl
3	37	52.9	566	4 US-09-491-522-7	Sequence 7, Appl
4	37	52.9	1205	4 US-09-491-522-11	Sequence 11, Appl
5	37	52.9	1211	4 US-09-491-522-5	Sequence 5, Appl
6	36	51.4	194	3 US-08-968-563-35	Sequence 35, Appl
7	36	51.4	194	4 US-08-968-563-35	Sequence 4, Appl
8	36	51.4	414	4 US-08-968-563-35	Sequence 4, Appl
9	36	51.4	414	4 US-08-968-563-35	Sequence 4, Appl
10	36	51.4	414	4 US-08-968-563-35	Sequence 4, Appl
11	35	50.0	803	4 US-09-613-192A-319	Sequence 319, Appl
12	35	50.0	395	4 US-09-613-192A-319	Sequence 10, Appl
13	35	50.0	395	4 US-09-613-192A-319	Sequence 10, Appl
14	35	50.0	516	4 US-08-948-564-12	Sequence 12, Appl
15	35	50.0	516	4 US-08-948-564-12	Sequence 12, Appl
16	35	50.0	686	2 US-09-462-844-2	Sequence 2, Appl
17	35	50.0	737	4 US-09-462-844-2	Sequence 2, Appl
18	34	48.6	737	4 US-09-462-844-2	Sequence 2, Appl
19	34	48.6	737	4 US-09-462-844-2	Sequence 2, Appl
20	34	48.6	296	4 US-09-653-270A-7	Sequence 7, Appl
21	34	48.6	296	4 US-09-653-270A-7	Sequence 7, Appl
22	34	48.6	296	4 US-09-653-270A-7	Sequence 7, Appl
23	34	48.6	403	4 US-09-503-391-8	Sequence 8, Appl
24	34	48.6	403	4 US-09-503-391-8	Sequence 8, Appl
25	34	48.6	766	4 US-07-754-918A-14	Sequence 14, Appl
26	34	48.6	766	4 US-07-754-918A-14	Sequence 14, Appl
27	34	48.6	1043	2 US-08-724-354D-4	Sequence 4, Appl

28	34	48.6	1043	3 US-09-270-984A-4	Sequence 4, Appl
29	34	48.6	1118	2 US-08-724-354D-2	Sequence 2, Appl
30	34	48.6	1118	3 US-09-270-984A-2	Sequence 2, Appl
31	33.5	47.9	1958	1 US-07-945-283-2	Sequence 9, Appl
32	33	47.1	28	1 US-08-427-072-9	Sequence 37, Appl
33	33	47.1	311	1 US-08-118-270-37	Sequence 37, Appl
34	33	47.1	311	5 PCT-0833-08528-37	Sequence 4857, Ap
35	33	47.1	348	4 US-09-134-001C-4857	Sequence 4, Appl
36	33	47.1	361	1 US-08-685-945B-4	Sequence 4, Appl
37	33	47.1	361	4 US-08-685-945B-4	Sequence 4, Appl
38	33	47.1	365	1 US-09-108-020-44	Sequence 2, Appl
39	33	47.1	365	1 US-08-390-162-2	Sequence 2, Appl
40	33	47.1	365	1 US-08-685-945B-2	Sequence 8, Appl
41	33	47.1	365	3 US-08-155-005A-8	Sequence 8, Appl
42	33	47.1	365	4 US-09-363-783-8	Sequence 4, Appl
43	33	47.1	378	1 US-08-416-756A-4	Sequence 4, Appl
44	33	47.1	378	4 US-08-880-965-4	Sequence 3, Appl
45	33	47.1	385	1 US-08-416-756A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-322-962-5
Sequence 5, Application US/08322962

Patent No. 546785

GENERAL INFORMATION:
APPLICANT: Detramond, Annick J
TITLE OF INVENTION: Tissue-Preferential Promoters
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,962
FILING DATE: 02-JUN-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/071,209
FILING DATE: 02-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/508,207
FILING DATE: 12-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: SPTULLI, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..75
OTHER INFORMATION: /note= "Protein product of pea
metallothionein reported by Evans et al."

US-08-322-962-5

Query Match

Best Local Similarity 100.0%; Score 70; DB 1; Length 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
DB 36 VIIIGVPAKIOFEG 49

RESULT 2

US-08-450-653-5
Sequence 5, Application US/08450653
Patent No. 6018099

GENERAL INFORMATION:

APPLICANT: Deffromond, Annick J
TITLE OF INVENTION: Tissue-Preferential Promoters
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,653
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/322,962
FILING DATE:
APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Spullin, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18039/CGC 1479/CONT2
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8615
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: Protein
LOCATION: 1-75
OTHER INFORMATION: /note="Protein product of pea
OTHER INFORMATION: metallothionein reported by Evans et al."
US-08-450-653-5

Query Match

Best Local Similarity 100.0%; Score 70; DB 3; Length 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
DB 36 VIIIGVPAKIOFEG 49

RESULT 3

US-09-491-522-7
Sequence 7, Application US/09491522

Patent No. 6428998

GENERAL INFORMATION:

APPLICANT: Collige, Alain
APPLICANT: Lapierre, Charles M.
TITLE OF INVENTION: PROCKOP, Darwin J.
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/886,333
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-7

Query Match

Best Local Similarity 52.9%; Score 37; DB 4; Length 566;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 VIIIGVPAKIOFEG 14
DB 270 VIIIGVDSVQFEG 283

RESULT 4

US-09-491-522-11
Sequence 11, Application US/09491522
Patent No. 6428998

GENERAL INFORMATION:

APPLICANT: Collige, Alain
APPLICANT: Lapierre, Charles M.
TITLE OF INVENTION: PROCKOP, Darwin J.
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

Mon Jan 13 10:35:08 2003

us-09-554-941-1.ra1

Page 3

OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/886,333
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1205 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-11
Query Match
Best local Similarity 52.9%; Score 37; DB 4; Length 1205;
Best local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 VILGVGPAKIOFEG 14
DB 264 VILGVDSVVOFEG 277
RESULT 5
US-09-491-522-5
Sequence 5, Application US/09491522
Patent No. 6428998
GENERAL INFORMATION:
APPLICANT: Collige, Alain
APPLICANT: Lapierre, Charles M.
APPLICANT: Prockop, Darwin J.
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/886,333
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 1211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-5
Query Match
Best local Similarity 52.9%; Score 37; DB 4; Length 1211;
Best local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 VILGVGPAKIOFEG 14
DB 270 VILGVDSVVOFEG 283
RESULT 6
US-08-968-563-35
Sequence 35, Application US/08968563
Patent No. 6013494
GENERAL INFORMATION:
APPLICANT: CHARLES E. NAKAMURA
APPLICANT: ANTHONY A. GATENBY
APPLICANT: AMY (KANG-HO) HSU
APPLICANT: RICHARD D. LA REAU
APPLICANT: SHARON L. HAYNIE
APPLICANT: MARIA DIAZ-TORRES
APPLICANT: DONALD E. TRIBBIE
APPLICANT: GREGORY M. WHITED
APPLICANT: VASANTHA NAGARAJAN
APPLICANT: MARK S. PAYNE
APPLICANT: STEPHEN K. PICATAGGIO
APPLICANT: RAMESCH V. NAIR
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANDIOL
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
STREET: 1870 SOUTH WINTON ROAD
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: 60/030,601
APPLICATION NUMBER: 13, 1996
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-773-0164
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: DHAB2
US-08-968-563-35

Query Match
Best Local Similarity 51.4%; Score 36; DB 3; Length 194;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VILGVGPA 8
DB 30 VVIGVGA 37

RESULT 7
US-08-969-683A-35
Sequence 35, Application US/08969683A
Patent No. 6136576
GENERAL INFORMATION:
APPLICANT: GENENCOR INTERNATIONAL, INC.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,683A
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/20873
FILING DATE: 13-NOV-1997
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Dedra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: DHAB2
US-08-969-683A-35

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 194;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VILGVGPA 8
DB 30 VVIGVGA 37

RESULT 8
US-09-608-917A-4
Sequence 4, Application US/09608917A
Patent No. 6409648
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
APPLICANT: Campisi, Judith
TITLE OF INVENTION: TRF1 Binding Protein
FILE REFERENCE: LBNL 18 1317
CURRENT APPLICATION NUMBER: US/09/608,917A
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 414
TYPE: PRT
ORGANISM: mouse
US-09-608-917A-4

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 414;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 4 GVGPAKIQP 12
DB 6 GVGPAKLP 14

RESULT 9
US-08-914-375C-72
Sequence 72, Application US/08914375C
Patent No. 6377893
GENERAL INFORMATION:
APPLICANT: Steven A. Benner
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: 1501 NW 68th Terrace
CITY: Gainesville
STATE: FL
COUNTRY: United States
ZIP: 32605-4147
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,375C
FILING DATE: 19-AUG-1997
CLASSIFICATION: 702/20
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352 392 7773
TELEFAX: 352 331 0462
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 459
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acid
ORIGINAL SOURCE:
ORGANISM: Trifolium repens
FEATURE:
OTHER INFORMATION: bgl2-trip non-cyanogenic strand-glucosylase precu
US-08-914-375C-72

Query Match
Best Local Similarity 51.4%; Score 36; DB 4; Length 459;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY	2	ILGVC	PAKIQ	EFG	14
			:		
Db	11	IFGAGSS	AYQ	EFG	23

1 GENERAL INFORMATION:
2 APPLICANT: PEUKERT, Karen; HAENEL, Frank; and EILERS,
3
4 APPLICANT: Martin
5 TITLE OF INVENTION: MYC-binding zinc finger proteins,
6 TITLE OF INVENTION: their preparation and their use
7 NUMBER OF SEQUENCES: 2
8 CORRESPONDENCE ADDRESSES:
9 ADDRESSEE: Kell & Weinkauff
10 STREET: 1101 Connecticut Avenue
11 City: Washington
12
13 STATE: D.C.
14
15 COUNTRY: USA
16
17 ZIP: 20036
18
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
21 COMPUTER: IBM AT-compatible, 80486 processor
22 OPERATING SYSTEM: MS-DOS version 6.1
23 SOFTWARE: Wordperfect version 8.0
24
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: US/09/063,035
27 FILING DATE: 21-Apr-1998
28
29 CLASSIFICATION: 514
30
31 INFORMATION FOR SEQ ID NO: 2:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 803 amino acids
34 TYPE: amino acid
35
36 TOPOLOGY: linear
37
38 US-09-063-035-2

51.4%; Score 36; DB 4: Length 803;
Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels

RESULT 11
 US-09-615-192A-319
 Sequence 319, Application US/09615192A
 Patent No. 6410718
 GENERAL INFORMATION:
 APPLICANT: Bloksberg, Leonard N.
 APPLICANT: Havukata, Ilka
 TITLE OF INVENTION: Materials and Methods for the
 TITLE OF INVENTION: Modification of Plant Lignin Content
 FILE REFERENCE: 11000.1003c4U US/09/615,192A
 CURRENT APPLICATION NUMBER: 2000-07-12
 CURRENT FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 08/975,316
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: US 08/7713,000
 PRIOR FILING DATE: 1996-09-11
 PRIOR APPLICATION NUMBER: US 09/169,789
 PRIOR FILING DATE: 1996-10-09
 NUMBER OF SEQ ID NOS: 405
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 319
 LENGTH: 141
 TYPE: PRT
 ORGANISM: Eucalyptus grandis
 US-09-615-192A-319

```
QY      1 VILGVGPAKIQF 12
        |||| | | : |
Db      71 VILGDNKAVVF 82
```

GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madhuri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Dan R.
APPLICANT: Waliron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
IS-09-036-987A-10

RESULT 13
US-09-370-700-10
; Sequence 10, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C

APPLICANT: Crawford, Kathryn P
 APPLICANT: Madduri, Krishnamurthy
 APPLICANT: Treadway, Patli J
 APPLICANT: Turner, Jan R
 TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 FILE REFERENCE: 50489 DIV1
 CURRENT APPLICATION NUMBER: US/09/370,700
 CURRENT FILING DATE: 1999-08-09
 EARLIER APPLICATION NUMBER: US 09/36987
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 10
 LENGTH: 395
 TYPE: PRT
 ORGANISM: Saccharopolyspora spinosa
 US-09-370-700-10

Query Match
 Best Local Similarity 50.0%; Score 35; DB 4; Length 395;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 LGVGPATQ 11
 DB 20 GAGPAPKQ 27

RESULT 14
 US-08-948-564-12
 Sequence 12, Application US/08948564
 Patent No. 6121512
 GENERAL INFORMATION:
 APPLICANT: Siminszky, Balazs
 APPLICANT: Dewey, Ralph E.
 APPLICANT: Corbin, Frederick T.
 TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
 TITLE OF INVENTION: Methods of Producing Herbicide-resistant Transgenic Plants
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESS: Virginia C. Bennett
 STREET: PO Box 37428
 CITY: Raleigh
 STATE: NO. 6121512el Carolina
 COUNTRY: USA
 ZIP: 27627
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/948,564
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Bennett, Virginia C.
 REGISTRATION NUMBER: 37,092
 REFERENCE/DOCKET NUMBER: 5051-409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-854-1400
 TELEFAX: 919-854-1401
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 516 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-948-564-12

Query Match
 Best Local Similarity 50.0%; Score 35; DB 3; Length 516;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LGVGPATQEEG 14
 DB 99 LVGGPAPKLSYDG 110
 RESULT 15
 US-09-201-641-6
 Sequence 6, Application US/09201641A
 Patent No. 6232530
 GENERAL INFORMATION:
 APPLICANT: Cunningham Jr, Francis X
 APPLICANT: Dellapenna, Dean
 TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
 FILE REFERENCE: Quest 41-162
 CURRENT APPLICATION NUMBER: US/09/201,641A
 CURRENT FILING DATE: 1998-11-30
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 6
 LENGTH: 516
 TYPE: PRT
 ORGANISM: Tagetes erecta
 US-09-201-641-6

Query Match
 Best Local Similarity 50.0%; Score 35; DB 4; Length 516;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VILGGPAPKIOFEG 14
 DB 98 VVIGGAPAGLAG 111

Search completed: January 13, 2003, 09:56:22
 Job time: 14.2051 secs

Title: US-09-554-941-1
 Perfect score: 70
 Sequence: 1 VILGVPAKIQFEG 14

Searched: 118974 seqs, 19401057 residues
118974 parameters:

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```

1: /cgn2_6/p/odata/1/pubppa/US08_NEW_PUB pep: *
2: /cgn2_6/p/odata/1/pubppa/PCT_NEW_PUB pep: *
3: /cgn2_6/p/odata/1/pubppa/US06_PUBCOMB pep: *
4: /cgn2_6/p/odata/1/pubppa/US07_PUBCOMB pep: *
5: /cgn2_6/p/odata/1/pubppa/US07_PUBCOMB pep: *
6: /cgn2_6/p/odata/1/pubppa/PCT05_PUBCOMB pep: *
7: /cgn2_6/p/odata/1/pubppa/US08_PUBCOMB pep: *
8: /cgn2_6/p/odata/1/pubppa/US09_NEW_PUB pep: *
9: /cgn2_6/p/odata/1/pubppa/US09_PUBCOMB pep: *
10: /cgn2_6/p/odata/1/pubppa/US10_NEW_PUB pep: *
11: /cgn2_6/p/odata/1/pubppa/US10_PUBCOMB pep: *
12: /cgn2_6/p/odata/1/pubppa/US06_NEW_PUB pep: *
13: /cgn2_6/p/odata/1/pubppa/US06_PUBCOMB pep: *
14: /cgn2_6/p/odata/1/pubppa/US06_PUBCOMB pep: *

```

14: /cgu24-27-7-0
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	55.7	28	10	US-09-864-761-36875	Sequence 36875, A
2	37	52.9	10	US-09-921-987B-5	Sequence 5, Appl1	
3	35	50.0	969	10	US-09-801-368-254	Sequence 254, Appl
4	35	50.0	366	10	US-09-738-626-6965	Sequence 6965, Ap
5	35	50.0	467	9	US-09-738-626-6965	Sequence 5777, Ap
6	35	50.0	510	9	US-09-323-998B-54	Sequence 2, Appl1
7	35	50.0	516	10	US-09-889-482-2	Sequence 458, Appl
8	35	50.0	737	9	US-09-939-980-458	Sequence 30, Appl
9	35	50.0	70	10	US-09-764-884-10	Sequence 279, Appl
10	34	48.6	121	9	US-09-764-887-279	Sequence 203, Appl
11	34	48.6	125	10	US-09-921-987B-5	Sequence 203, Appl
12	34	48.6	135	10	US-09-921-217-203	Sequence 203, Appl
13	34	48.6	135	10	US-09-833-263-203	Sequence 45, Appl1
14	34	48.6	135	10	US-09-978-295A-45	Sequence 45, Appl1
15	34	48.6	359	9	US-09-978-697-45	Sequence 45, Appl1
16	34	48.6	359	9	US-09-978-932A-45	Sequence 45, Appl1
17	34	48.6	359	9	US-09-599-832A-45	Sequence 45, Appl1
18	34	48.6	359	9	US-09-978-189-45	Sequence 45, Appl1
19	34	48.6	359	12	US-10-052-586-42	Sequence 42, Appl

45	32	45.7	58	10	US-09-1764-877-1709	Sequence 1709, Ap
44	32	45.7	32	10	US-09-1764-877-1709	Sequence 1709, Ap
43	32	45.7	32	10	US-09-1764-877-1709	Sequence 1709, Ap
42	32	45.7	32	10	US-09-1764-877-1709	Sequence 1709, Ap
41	33	47.1	780	9	US-09-864-761-39756	Sequence 39756, Ap
40	33	47.1	481	9	US-09-815-242-5688	Sequence 4706, Ap
39	33	47.1	471	10	US-09-815-242-5688	Sequence 5688, Ap
38	33	47.1	467	10	US-09-815-242-10943	Sequence 8, Appl
37	33	47.1	467	10	US-09-735-933-6	Sequence 10943, A
36	33	47.1	466	10	US-09-735-933-6	Sequence 6, Appl
35	33	47.1	466	10	US-09-735-933-6	Sequence 6, Appl
34	33	47.1	466	10	US-09-815-242-14017	Sequence 5, Appl
33	33	47.1	448	10	US-09-815-242-14017	Sequence 14017, A
32	33	47.1	442	10	US-09-815-242-13272	Sequence 10021, A
31	33	47.1	442	10	US-09-815-242-13272	Sequence 13272, A
30	33	47.1	233	10	US-09-815-242-11817	Sequence 11817, A
29	33	47.1	233	10	US-09-815-242-11817	Sequence 11817, A
28	33	47.1	37	10	US-09-925-301-1010	Sequence 1010, Ap
27	34	48.6	811	9	US-09-745-765-82	Sequence 82, Appl
26	34	48.6	811	9	US-09-745-765-82	Sequence 82, Appl
25	34	48.6	678	9	US-09-738-626-5817	Sequence 5817, Ap
24	34	48.6	500	9	US-09-738-626-5817	Sequence 5817, Ap
23	34	48.6	500	9	US-09-738-626-5817	Sequence 5817, Ap
22	34	48.6	500	9	US-09-738-626-5817	Sequence 5817, Ap
21	34	48.6	478	10	US-09-815-242-11208	Sequence 6755, Ap
20	34	48.6	475	10	US-09-815-242-11208	Sequence 4296, Ap
19	34	48.6	475	10	US-09-815-242-11208	Sequence 6393, Ap
18	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
17	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
16	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
15	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
14	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
13	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
12	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
11	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
10	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
9	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
8	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
7	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
6	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
5	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
4	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
3	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
2	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap
1	34	48.6	475	10	US-09-815-242-11208	Sequence 4369, Ap

ALIGNMENTS

RESULT 1
 US-09-864-761-36875 Application US/09864761
 Sequence 36875, Appication US/09864763A1
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wenshang
 APPLICANT: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 TITLE OF INVENTION: Aecm1c-X-1
 FILE REFERENCE: Aecm1c-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,667
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36875
LENGTH: 28
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL049539.19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
US-09-864-761-36875
EST_HUMAN HIT: AV659818.1, EVALUATE 4.00e-08

```

```

Query Match
Best Local Similarity 55.7%; Score 39; DB 10; Length 28;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 ILGVGPAKIQFE 13
DB 7 VLAAGPQLOAE 18

```

```

RESULT 2
US-09-321-987B-5
Sequence 5, Application US/09321987B
Patent No. US20020102210A1
GENERAL INFORMATION:
APPLICANT: Kimble, Judith E
TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
FILE REFERENCE: 960296.95386
CURRENT APPLICATION NUMBER: US/09/321,987B
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,170
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/129,023
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 969
TYPE: PRT
ORGANISM: Bovine
US-09-321-987B-5

```

```

Query Match
Best Local Similarity 52.9%; Score 37; DB 10; Length 969;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 VILGVGPAKIQFE 14
DB 264 VILGVDDSVQDFHG 277

```

```

RESULT 3
US-09-801-368-254
Sequence 254, Application US/09801368

```

```

Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 254
LENGTH: 366
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-254

```

```

Query Match
Best Local Similarity 50.0%; Score 35; DB 10; Length 366;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 2 ILGVGPAKIQFE 14
DB 116 VMAVMPPEVQDFG 128

```

```

RESULT 4
US-09-738-626-6965
Sequence 6965, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAKIO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6965
LENGTH: 467
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6965

```

Query Match 50.0%; Score 35; DB 9; Length 467;
 Best Local Similarity 35.7%; Pred. No. 84;
 Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 VILGVPAKIOPEG 14
 :||| || :|
 Db 172 VVVGAGPTGVELAG 185

RESULT 5

US-09-738-626-5777
 ; Sequence 5777, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEGA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 5777
 ; LENGTH: 510
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-5777

Query Match 50.0%; Score 35; DB 9; Length 510;
 Best Local Similarity 50.0%; Pred. No. 93;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 VILGVPAKIOPEG 14
 :||| || :|
 Db 260 IILGISPALDAVEG 273

RESULT 6

US-09-323-998D-54
 ; Sequence 54, Application US/09323998D
 ; Patent No. US20020102631A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; APPLICANT: SUN, ZAIREN
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
 ; FILE REFERENCE: 108172-09019
 ; CURRENT APPLICATION NUMBER: US/09/323,998D
 ; CURRENT FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 09/086,724
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 09/086,725
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 54
 ; LENGTH: 516
 ; TYPE: PRT
 ; ORGANISM: Tagetes erecta
 ; US-09-323-998D-54

Query Match 50.0%; Score 35; DB 10; Length 516;
 Best Local Similarity 42.9%; Pred. No. 95;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 VILGVPAKIOPEG 14
 :||| || :|
 Db 98 VVIGCGPAGLALAG 111

RESULT 7

US-09-899-482-2
 ; Sequence 2, Application US/09899482
 ; Patent No. US20020006641A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Quax, Wilhelmus J.
 ; TITLE OF INVENTION: Increasing Production of Proteins in Microorganisms
 ; FILE REFERENCE: GC385-PCT
 ; CURRENT APPLICATION NUMBER: US/09/899,482
 ; CURRENT FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: EP 97305286.3
 ; PRIOR FILING DATE: 1997-07-16
 ; PRIOR APPLICATION NUMBER: EP 97305344.0
 ; PRIOR FILING DATE: 1997-07-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 737
 ; TYPE: PRT
 ; ORGANISM: Bacillus Subtilis
 ; US-09-899-482-2

Query Match 50.0%; Score 35; DB 10; Length 737;
 Best Local Similarity 53.8%; Pred. No. 1,4e+02;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 IILGVPAKIOPEG 14
 :||| || :|
 Db 74 VILGVSEPIQIEG 86

RESULT 8

US-09-939-980-458
 ; Sequence 458, Application US/09939980
 ; Patent No. US20020082234A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Black, Michael
 ; Burnham, Martin
 ; Hodgson, John
 ; Knowles, David
 ; Lonetto, Michael
 ; Nicholas, Richard
 ; Pratt, Julie
 ; Reichard, Richard
 ; Rosenberg, Martin
 ; Ward, Judith
 ; TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
 ; Polypeptides and Their Uses
 ; NUMBER OF SEQUENCES: 534
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette

```

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm1, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 458:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 70 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 458:
US-09-939-980-458

Query Match      48.6%: Score 34; DB 10; Length 70;
Best Local Similarity 70.0%: Pred. No. 13;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GVGPAKIQFE 13
    ||| | |
Db 42 GVGPKENFE 51

RESULT 9
US-09-764-884-30
; Sequence 30, Application US/09764884
; Patent No. US20020161208A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT03
; CURRENT APPLICATION NUMBER: US/09/764,884
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-884-30

Query Match      48.6%: Score 34; DB 9; Length 121;
Best Local Similarity 38.5%: Pred. No. 26;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 ILGVGPAKIQFEG 14
    :|:| | |
Db 96 VIGAGKIVFDG 108

RESULT 10
US-09-764-887-279
; Sequence 279, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAl13

```

```

; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 279
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-887-279

Query Match      48.6%: Score 34; DB 10; Length 125;
Best Local Similarity 50.0%: Pred. No. 27;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LGVGPAKIQFEG 14
    ||:| | |
Db 92 LGMGPKAEPQG 103

RESULT 11
US-10-025-380-203
; Sequence 203, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-203

Query Match      48.6%: Score 34; DB 9; Length 135;
Best Local Similarity 54.5%: Pred. No. 29;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 GVGPAKIQFEG 14
    || | | |
Db 80 GVNPGKIDVEG 90

RESULT 12
US-09-922-217-203
; Sequence 203, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy

```

```

; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-203

```

```

Query Match      48.6%; Score 34; DB 10; Length 135;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 GVGPAKIQFEG 14
   |||:|:|
Db 80 GVGPKLIDVEG 90

```

```

RESULT 13
US-09-833-263-203
; Sequence 203, Application US/09833263
; Patent No. US2002010547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR COLON CANCER AND METHODS FOR THEIR USE
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 203
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-203

```

```

Query Match      48.6%; Score 34; DB 10; Length 135;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 GVGPAKIQFEG 14
   |||:|:|
Db 80 GVGPKLIDVEG 90

```

```

RESULT 14
US-09-978-295A-45
; Sequence 45, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Flivarov, Ellen
; APPLICANT: Fong, Sherman

```

```

; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/07450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/07632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105

```

PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080107
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080194
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/080327
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080328
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080333
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080334
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/081070
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081049
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081071
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081955
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081819
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081952
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082804
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 48.6%; Score 34; DB 9; Length 359;
 Best Local Similarity 77.8%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GPAKIOFEG 14
 Db 84 GPAKIOFEG 92

RESULT 15
 US-09-978-697-45
 Sequence 45, Application US/09978697
 Patent No. US20020169284A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Baker Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan
 APPLICANT: Ferrari, Napoleon
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerlitsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT FILING DATE: 2001-10-16
CURRENT FILING DATE: 2001-09/978,697
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441

```

; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match 48.6%; Score 34; DB 9; Length 359;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 6 GPAKQFEG 14
    |||||
Db 84 GPAKPFEG 92

```

Search completed: January 13, 2003, 10:02:49
Job time : 65.9744 secs

Mon Jan 13 10:35:11 2003

us-09-554-941-11.rapb

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:54:16 ; Search time 64.9744 Seconds
(without alignments)
4.180 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 GASTRGRTAPQ 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 segs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications-AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	56.2	49	10	US-09-864-761-47113
2	39	53.4	295	10	US-09-893-737-324
3	39	53.4	423	10	US-09-729-674-160
4	39	53.4	428	10	US-09-780-016-22
5	39	53.4	453	10	US-09-780-016-18
6	39	53.4	484	10	US-09-780-016-26
7	39	53.4	507	10	US-09-780-016-2
8	39	53.4	509	10	US-09-780-016-24
9	39	53.4	532	10	US-09-780-016-20
10	39	53.4	70	10	US-09-864-761-39451
11	36	49.3	101	9	US-09-989-920-184
12	36	49.3	330	9	US-09-794-591-2
13	36	49.3	374	9	US-09-995-225-6
14	36	49.3	407	9	US-09-738-626-5433
15	35.5	48.6	639	10	US-09-782-906-2
16	35.5	48.6	639	10	US-09-782-906-3
17	35.5	48.6	639	10	US-09-782-906-4
18	35.5	48.6	639	10	US-09-782-906-5
19	35	47.9	86	10	US-09-987-108-30

20	35	47.9	111	10	US-09-764-869-848	Sequence 848, App
21	35	47.9	460	10	US-09-925-297-584	Sequence 584, App
22	35	47.9	1070	10	US-09-735-367B-6	Sequence 6, Appl1
23	35	47.9	2005	10	US-09-735-367B-3	Sequence 3, Appl1
24	35	47.9	2063	10	US-09-735-367B-2	Sequence 2, Appl1
25	34.5	47.3	651	10	US-09-837-992-3	Sequence 3, Appl1
26	34	46.6	147	9	US-09-922-199A-8	Sequence 8, Appl1
27	34	46.6	185	9	US-09-922-199A-7	Sequence 7, Appl1
28	34	46.6	203	10	US-09-815-242-10084	Sequence 10084, A
29	34	46.6	286	10	US-09-911-826A-4	Sequence 4, Appl1
30	34	46.6	309	9	US-09-976-059-34	Sequence 34, Appl1
31	34	46.6	1298	9	US-09-825-288A-2	Sequence 2, Appl1
32	33.5	45.9	1499	10	US-09-911-826A-2	Sequence 115, App
33	33	45.2	288	10	US-09-888-615-115	Sequence 48476, A
34	33	45.2	31	10	US-09-864-761-48476	Sequence 48910, A
35	33	45.2	34	10	US-09-864-761-48910	Sequence 40180, A
36	33	45.2	68	10	US-09-864-761-40180	Sequence 35131, A
37	33	45.2	127	10	US-09-864-761-35131	Sequence 38443, A
38	33	45.2	141	10	US-09-864-761-38443	Sequence 20, Appl1
39	33	45.2	155	9	US-10-014-717-20	Sequence 412, App
40	33	45.2	236	10	US-09-731-872-412	Sequence 1075, App
41	33	45.2	244	10	US-09-764-864-1075	Sequence 310, App
42	33	45.2	276	10	US-09-801-368-310	Sequence 434, App
43	33	45.2	280	10	US-09-925-302-454	Sequence 56, Appl1
44	33	45.2	283	10	US-09-920-552-56	Sequence 10817, A
45	33	45.2	300	10	US-09-815-242-10817	

ALIGNMENTS

RESULT 1

US-09-864-761-47113

Sequence 47113, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, David R.

APPLICANT: Rank, David K.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Weisheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeonica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47113
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005384.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
; OTHER INFORMATION: SWISSPROT HIT: Q04855, EVALU 2.10e+00
US-09-864-761-47113

```

```

Query Match
Best Local Similarity 56.2%; Score 41; DB 10; Length 49;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYAPOP 13
DB 20 VEGHRTVPR 29

```

```

RESULT 2
; US-09-893-737-324
; Sequence 324, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 324
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-324

```

```

Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 295;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 RYTAPOPO 14
DB 186 KYTAPOPE 193

```

```

RESULT 3
; US-09-729-674-160
; Sequence 160, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl

```

```

; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 160
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-160

```

```

Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 423;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 7 RYTAPOPO 14
DB 88 KYTAPOPE 95

```

```

RESULT 4
; US-09-780-016-22
; Sequence 22, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020004591A1 Human Proteases and
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,294
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 428
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-22

```

```

Query Match
Best Local Similarity 54.5%; Score 39; DB 10; Length 428;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 IEGRYAPOP 14
DB 270 VNGRFTAPOAE 280

```

```

RESULT 5
; US-09-780-016-18
; Sequence 18, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 453
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-18

```

```

Query Match      53.4%; Score 39; DB 10; Length 453;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 4 IEGRYTAPOQ 14
: ||:||||:
Db 295 VNGRFTAPOAE 305

```

```

RESULT 6
US-09-780-016-26
; Sequence 26, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 484
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-26

```

```

Query Match      53.4%; Score 39; DB 10; Length 484;
Best Local Similarity 54.5%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 4 IEGRYTAPOQ 14
: ||:||||:
Db 326 VNGRFTAPOAE 336

```

```

RESULT 7
US-09-780-016-2
; Sequence 2, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory

```

```

; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 507
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-2

```

```

Query Match      53.4%; Score 39; DB 10; Length 507;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 4 IEGRYTAPOQ 14
: ||:||||:
Db 349 VNGRFTAPOAE 359

```

```

RESULT 8
US-09-780-016-24
; Sequence 24, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0132-USA
; CURRENT APPLICATION NUMBER: US/09/780,016
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 509
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-780-016-24

```

```

Query Match      53.4%; Score 39; DB 10; Length 509;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 4 IEGRYTAPOQ 14
: ||:||||:
Db 351 VNGRFTAPOAE 361

```

```

RESULT 9
US-09-780-016-20
; Sequence 20, Application US/09780016
; Patent No. US20020004591A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John

```

```
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abulin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20020004591A1 Human Proteases and
FILE REFERENCE: LEX-0132-USA
CURRENT APPLICATION NUMBER: US 09/780,016
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,294
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 20
LENGTH: 532
TYPE: PRT
ORGANISM: homo sapiens
US-09-780-016-20
```

```
Query Match
Best Local Similarity 53.4%; Score 39; DB 10; Length 532;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 IEGRYTAPQ 14
DB 374 VNGRTAPQAE 384
```

```
RESULT 10
US-09-864-761-39451
Sequence 39451, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
```

```
PRIOR APPLICATION NUMBER: US 60/234,667
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39451
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006075.1
OTHER INFORMATION: EXPRESSED IN HETA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
OTHER INFORMATION: SWISSPROT HIT: Q09457, EVALUATE 5.00e-03
OTHER INFORMATION: EST_HUMAN HIT: A0118435.1, EVALUATE 8.00e-37
US-09-864-761-39451
```

```
Query Match
Best Local Similarity 49.3%; Score 36; DB 10; Length 70;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 QASIEGRYAPQ 13
DB 14 QNCIPATRYTAPHP 26
```

```
RESULT 11
US-09-989-920-184
Sequence 184, Application US/09989920
Patent No. US20020172957A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Hervé
APPLICANT: Chen, Sel-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US 09/989,920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 184
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-184
```

```
Query Match
Best Local Similarity 49.3%; Score 36; DB 9; Length 101;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 6 GRYTAPQ 13
DB 4 GRYQAPRP 11
```

```
RESULT 12
US-09-794-591-2
Sequence 2, Application US/09794591
Patent No. US20010018198A1
GENERAL INFORMATION:
```

```
;; APPLICANT: Pulst, Stefan M.
;; APPLICANT: Shidata, Hiroki
;; TITLE OF INVENTION: Nucleic Acids Encoding Ataxin-2 Binding Proteins,
;; TITLE OF INVENTION: Products Related Thereto and Methods of Using Same
;; FILE REFERENCE: CE 3093
;; CURRENT FILING DATE: 2001-02-26
;; PRIOR APPLICATION NUMBER: 09/145,391
;; PRIOR FILING DATE: 1998-09-01
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 330
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-591-2

Query Match          49.3%; Score 36; DB 10; Length 330;
Best Local Similarity 53.8%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 QASIEGRYTAPOP 13
        | | | | |
Db      34 QNGIPAEYTAAPH 46

RESULT 13
US-09-995-225-6
;; Sequence 6, Application US/09995225
;; Publication No. US20020193584A1
GENERAL INFORMATION:
;; APPLICANT: Chen, Ruoping
;; APPLICANT: Chu, Zhi Liang
;; APPLICANT: Dang, Huang T.
;; APPLICANT: Lowitz, Kevin P.
;; APPLICANT: Pride, Cameron
;; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
;; TITLE OF INVENTION: Receptors
;; FILE REFERENCE: AREN-0308
;; CURRENT APPLICATION NUMBER: US/09/995,225
;; CURRENT FILING DATE: 2001-11-26
;; PRIOR APPLICATION NUMBER: 09/170,496
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: PCT/US99/23938
;; PRIOR FILING DATE: 1998-10-13
;; PRIOR APPLICATION NUMBER: 60/253,404
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/255,366
;; PRIOR FILING DATE: 2000-12-12
;; PRIOR APPLICATION NUMBER: 60/270,286
;; PRIOR FILING DATE: 2001-02-20
;; PRIOR APPLICATION NUMBER: 60/282,365
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/270,266
;; PRIOR FILING DATE: 2001-02-20
;; PRIOR APPLICATION NUMBER: 60/282,032
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/282,358
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/282,356
;; PRIOR FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/290,917
;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: 60/309,208
;; PRIOR FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6
;; LENGTH: 374
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
OTHER INFORMATION: No. US20020193584A1el Sequence
```

```
US-09-995-225-6

Query Match          49.3%; Score 36; DB 9; Length 374;
Best Local Similarity 63.6%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      4 IEGRYTAPOPQ 14
        | | | | |
Db      26 LEGRETAQOPR 36

RESULT 14
US-09-738-626-5433
;; Sequence 5433, Application US/09738626
;; Publication No. US20020197605A1
GENERAL INFORMATION:
;; APPLICANT: NAKAGAWA, SATOSHI
;; APPLICANT: MIZOGUCHI, HIROSHI
;; APPLICANT: ANDO, SEIKO
;; APPLICANT: HAYASHI, MIKIRO
;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAOKO
;; APPLICANT: SENOH, AKIHIRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OZAKI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
;; CURRENT FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: JP 99/377484
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: JP 00/159162
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: JP 00/280988
;; PRIOR FILING DATE: 2000-08-03
;; NUMBER OF SEQ ID NOS: 7059
;; SOFTWARE: PatentIn ver. 3.0
;; SEQ ID NO 5433
;; LENGTH: 407
;; TYPE: PRT
;; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5433

Query Match          49.3%; Score 36; DB 9; Length 407;
Best Local Similarity 53.8%; Pred. No. 79;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 ASIEGRYTAPOPQ 14
        | | | | |
Db      248 AAAPQOYQAPOPQ 260

RESULT 15
US-09-782-906-2
;; Sequence 2, Application US/09782906
;; Patent No. US20010051369A1
GENERAL INFORMATION:
;; APPLICANT: Delagrave, Simon
;; APPLICANT: Rittenhouse Pruss, Jennifer L.
;; APPLICANT: Murphy, Dennis J.
;; APPLICANT: Maffia III, Anthony M.
;; APPLICANT: Bylina, Edward J.
;; APPLICANT: Coleman, William J.
;; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Met
;; FILE REFERENCE: HER-0040
;; CURRENT APPLICATION NUMBER: US/09/782,906
;; CURRENT FILING DATE: 2001-02-14
;; PRIOR APPLICATION NUMBER: 60/185,001
;; PRIOR FILING DATE: 2000-02-25
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.0
```

; SEQ ID NO 2
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Dactylium dendroides
US-09-782-906-2

Query Match 48.6%; Score 35.5; DB 10; Length 639;
Best Local Similarity 69.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY 1 QASIEGRYTAPOP 13
Db 147 QAS---SYTAPOP 156

Search completed: January 13, 2003, 10:02:50
Job time : 65.9744 secs

Mon Jan 13 10:35:11 2003

us-09-554-941-11.rpr

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 12.9231 seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 QASIEGRYTAPOPO 14

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	61.6	326	2 A59232	ABA-responsive pro
2	43	58.9	295	4 JC6530	laminin receptor p
3	42	57.5	424	2 S76359	hypothetical prote
4	41	56.2	224	2 B75388	NADH dehydrogenase
5	40	54.8	1263	2 AH2011	heterocyst glycoli
6	39	53.4	362	2 S78515	single-stranded nu
7	39	53.4	365	2 S42471	hnRNP protein E2 -
8	39	53.4	506	2 T35923	probable tRNA synt
9	39	53.4	508	2 A33378	fasciclin III prec
10	39	53.4	649	2 T00062	protein F17F8.23 l
11	39	53.4	1571	2 T24965	hypothetical prote
12	38	52.1	271	2 T24965	hypothetical prote
13	38	52.1	293	2 AC3464	vacuolar sorting r
14	38	52.1	628	2 T02602	transposase - phag
15	38	52.1	628	2 T02604	Bassoon protein -
16	38	52.1	662	2 T42730	glycerophosphoryl
17	38	52.1	3942	1 T42730	diuretic hormone p
18	38	52.1	786	2 D75630	hypothetical prote
19	37.5	51.4	138	1 SWMOD	probable pili-rel
20	37	50.7	156	2 T29730	probable right ori
21	37	50.7	286	2 AF0610	xyloglucan endo-1,
22	37	50.7	288	2 AE0056	hypothetical prote
23	37	50.7	296	2 C49539	multidrug-efflux t
24	37	50.7	396	2 D83520	hypothetical prote
25	37	50.7	401	2 F69634	hypothetical prote
26	37	50.7	478	2 T51274	hypothetical prote
27	37	50.7	608	2 T02299	Spot 3 protein and
28	37	50.7	623	2 T47542	

30	37	50.7	623	2 T06794	vacuolar sorting r
31	37	50.7	624	2 T00044	vacuolar sorting r
32	37	50.7	625	2 F84706	probable vacuolar
33	37	50.7	626	2 T04895	vacuolar sorting r
34	37	50.7	768	2 I53821	P-selectin - rat
35	37	50.7	768	2 A42755	P-selectin precurs
36	37	50.7	798	2 S20881	homeotic protein p
37	37	50.7	961	2 JC2200	alpha-mannosidase
38	37	50.7	1116	2 D97695	COBN protein homol
39	37	50.7	1116	2 AC2921	copalamin biosynth
40	37	50.7	1124	2 T30340	dsRNA adenosine de
41	37	50.7	2219	2 T27684	hypothetical prote
42	37	50.7	4385	2 T29042	hypothetical prote
43	36.5	50.0	337	2 T32142	hypothetical prote
44	36	49.3	22	2 I51565	protein-tyrosine k
45	36	49.3	176	2 AE0526	2'-5' RNA ligase (

ALIGNMENTS

RESULT 1
A59232
ABA-responsive protein - barley
C:Species: Hordeum vulgare (barley)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text-change 02-Jun-2000
C:Accession: A59232
R:Jili, J.H.; Luo, M.; Mohapatra, S.S.
submitted to GenBank, September 1997
A:Reference number: A59232
A:Accession: A59232
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-326 <LNU>
A:Cross-references: GB:AF026538; NID:94103634; PIDN:AAD09343.1; PID:94103635
A:Experimental source: dev stage embryo
C:Superfamily: barley ABA-responsive protein

Query Match 61.6%; Score 45; DB 2; Length 326;
Best Local Similarity 70.0%; Pred. No. 2.3;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 IEGRYTAPOPO 13
DB 76 IEGRYTAPOPO 85

RESULT 2
JC6530
laminin receptor processed pseudogene LAMRL5 - human
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1998 #sequence_revision 28-Aug-1998 #text-change 28-Aug-1998
C:Accession: JC6530
R:Richardson, M.P.; Braybrook, C.; Tham, M.; Moore, G.E.; Stanier, P.
Gene 206, 145-150, 1998
A:Title: Molecular cloning and characterization of a highly conserved human 67-kDa la
A:Reference number: JC6530; MUID:98121324; PMID:9461426
A:Accession: JC6530
A>Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-295 <RIC>
A:Experimental source: Brain
C:Comment: No evidence could be found that this intronless gene sequence is expressed

C:Gene: LAMRL5
A:Gene position: Xq21.3
A:Map position: status absent
A:Introns: status absent
C:Keywords: brain; glycoprotein; laminin binding; pseudogene; receptor

Query Match 58.9%; Score 43; DB 4; Length 295;
Best Local Similarity 63.6%; Pred. No. 4.7;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 IEGRYTAPOPO 14
 Db 227 LEGOWTAPAPQ 237

RESULT 3

S76359
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.
 A:Variate: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S76359
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yasuda, M.
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S76359
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-424 <KAN>
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:91001484; PIDN:BA110211.1; PID:9100158
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: site-specific methyltransferase (cytosine-specific) EcoRII

Query Match

Best Local Similarity 57.5%; Score 42; DB 2; Length 424;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASIEGRTAPOP 13
 Db 319 ASDGATAPAP 330

RESULT 4

B75388
 NADH dehydrogenase I, C subunit - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75388
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A/5250; MUID:20036896; PMID:10567266
 A:Accession: B75388
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <WHI>
 A:Cross-references: GB:AE001994; GB:AE000513; NID:96459259; PIDN:AAF11071.1; PID:9645926
 C:Genetics:
 A:Gene: DR1504
 A:Map position: 1

Query Match

Best Local Similarity 56.2%; Score 41; DB 2; Length 224;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 RYTAPOPO 14
 Db 80 RYTAPOPO 87

RESULT 5

AN2011
 heterocyst glycolipid synthase [Imported] - *Nostoc* sp. (strain PCC 7120)
 C:Species: *Nostoc* sp.
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AN2011

R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AN2011
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1263 <KAN>
 A:Cross-references: GB:BA000019; PIDN:BA878012.1; PID:911135466; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: hglE

Query Match

Best Local Similarity 54.8%; Score 40; DB 2; Length 1263;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEGRYTAPOPO 14
 Db 1154 LEPRYSAPAPQ 1164

RESULT 6

S78515
 single-stranded nucleic acid-binding protein CBP - mouse
 N:Alternate names: hnRNP protein X
 C:Species: *Mus musculus* (house mouse)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999
 C:Accession: S78515; S45261; S78514; S41652; I48281
 R:Hahn, K.B.; Kim, G.; Turch, C.; Smale, S.T.
 submitted to the EMBL Data Library, June 1993
 A:Reference number: S78515
 A:Accession: S78515
 A:Molecule type: mRNA
 A:Residues: 1-362 <HAN>
 A:Cross-references: EMBL:L19661; NID:9436893; PIDN:AAA03705.1; PID:9436894
 R:Goller, M.; Funke, B.; Gene-Baecker, C.; Kroege, B.; Lottspeich, F.; Horak, I.
 Nucleic Acids Res. 22, 1885-1889, 1994
 A:Title: Murine protein which binds preferentially to oligo-C-rich single-stranded nu
 A:Reference number: I48281; MUID:94268912; PMID:8208614
 A:Accession: S45261
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-193, 225-300, 302-362 <GOL>
 A:Cross-references: EMBL:X75947; NID:9495127; PIDN:CAA53546.1; PID:9495128
 A:Accession: S78514
 A:Molecule type: Protein
 A:Residues: 12-39; 172-193; 225-229; 306-322 <GOW>
 R:Hahn, K.; Kim, G.; Turch, C.W.; Smale, S.T.
 Nucleic Acids Res. 21, 3894, 1993
 A:Title: Isolation of a murine gene encoding a nucleic acid-binding protein with homo
 A:Reference number: S41652; MUID:93376518; PMID:8367306
 A:Accession: S41652
 A:Molecule type: mRNA
 A:Residues: 16-51; 100-135; 287-322 <HAM>
 C:Genetics:
 A:Gene: CBP
 C:Keywords: alternative splicing; RNA binding; single-stranded DNA binding
 F:1-362/Product: single-stranded nucleic acid-binding protein CBP, long form #status
 F:1-193, 225-362/Product: single-stranded nucleic acid-binding protein CBP, short form

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 362;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SIEGRTAPOP 13
 Db 227 TIOGRTAPAPQ 237

RESULT 7

S42471

hnRNP protein e2 - human

N.Alternate names: nucleic acid-binding protein; protein PCBP-2

C.Species: Homo sapiens (man)

C.Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C.Accession: S65679; S42471

R.Leffers, H.; Dejgaard, K.; Celis, J.E.

Eur. J. Biochem. 230, 447-453, 1995

A.Title: Characterisation of two major cellular poly(rC)-binding human proteins, each co

A.Reference number: S65678; MUID:9333178; PMID:7607214

A.Accession: S65679

A.Status: nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-365 <LEP>

A.Cross-references: EMBL:X78136; NID:q460772; PIDN:CAA55015.1; PID:q460773

A.Experimental source: AMA cells (transformed human amnion cells)

A.Note: Submitted to the EMBL Data Library, March 1994

C.Keywords: RNA binding

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 365;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SIEGRYTAPOP 13

Db 231 TIQGVYAIPOP 241

RESULT 8

probable tRNA synthetase - Streptomyces coelicolor

C.Species: Streptomyces coelicolor

C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C.Accession: T35923

R.Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1999

A.Reference number: Z21551

A.Accession: T35923

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-506 <SEP>

A.Cross-references: EMBL:AL035206; PIDN:CAA22745.1; GSPDB:GN00070; SCQEDB:SC9B5.03

A.Experimental source: strain A3(2)

C.Genetics:

A.Gene: SCQEDB:SC9B5.03

C.Superfamily: methionine-tRNA ligase

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 506;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ASTEGRYTAPOP 13

Db 481 AAVEARYTAAP 492

RESULT 9

faciclin III precursor - fruit fly (Drosophila melanogaster)

C.Species: Drosophila melanogaster

C.Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 24-Sep-1998

C.Accession: A33378

R.Snow, P.M.; Bieber, A.J.; Goodman, C.S.

Cell 59, 313-323, 1989

A.Title: Faciclin III: a novel homophilic adhesion molecule in Drosophila.

A.Reference number: A33378; MUID:90030406; PMID:2509076

A.Accession: A33378

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-508 <SNO>

A.Cross-references: GB:M27813; NID:q157423; PID:q157424

C.Genetics:

A.Gene: FlyBase:Pas3

A.Cross-references: FlyBase:FBgn0000636

C.Keywords: phosphoprotein; transmembrane protein

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 508;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 14

Db 279 QGRIDGRISATPEPQ 292

RESULT 10

G86434

protein F17F8.23 [imported] - Arabidopsis thaliana

C.Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C.Accession: G86434

R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A.Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A.Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A.Reference number: A86141; MUID:21016719; PMID:11130712

A.Accession: G86434

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-649 <STPO>

A.Cross-references: GB:AE005172; NID:q9755389; PIDN:AAF98196.1; GSPDB:GN00141

C.Genetics:

A.Gene: F17F8.23

A.Map position: 1

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 649;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 EGRYTAPOPQ 14

Db 255 QGRYCAPDPE 264

RESULT 11

hypothetical protein KIAA0434 - human (fragment)

C.Species: Homo sapiens (man)

C.Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 23-Mar-2001

C.Accession: T00062

R.Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka,

submitted to the EMBL Data Library, October 1997

A.Description: Prediction of the coding sequences of unidentified human genes. VIII.

A.Reference number: Z14082

A.Accession: T00062

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-1571 <ISH>

A.Cross-references: EMBL:AB007894; NID:q2662148; PIDN:BAA23707.1; PID:q2662149

A.Experimental source: brain; clone HH2165

C.Genetics:

A.Note: KIAA0434

Query Match

Best Local Similarity 53.4%; Score 39; DB 2; Length 1571;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOPQ 14

Db 1389 QPOLQGRQAAPEPQ 1402

```

RESULT 12
C81322
probable periplasmic protein Cj1169c [imported] - Campylobacter jejuni (strain NCTC 1116
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: C81322
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A:Reference number: AB1250; MUID:20150912; PMID:10686204
A:Accession: C81322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:g6568444; PIDN:CAB73423.1; PID:g656860
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1169c

Query Match          52.1%; Score 38; DB 2; Length 75;
Best Local Similarity 46.2%; Pred. No. 8.6;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 ASIEGRTAPOPQ 14
||: ||: ||: ||:
Db 19 ASIMAKYEAPPE 31

RESULT 13
T24965
hypothetical protein T18D3.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T24965
R:McMurray, A.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19962
A:Accession: T24965
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-271 <MIL>
A:Cross-references: EMBL:Z68119; PIDN:CAA92195.1; GSPDB:GN00028; CESP:T18D3.6
A:Experimental source: clone T18D3
C:Genetics:
A:Gene: CESP:T18D3.6
A:Map position: X
A:Introns: 247/3

Query Match          52.1%; Score 38; DB 2; Length 271;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
||||| |||
Db 33 EGRYTPPLPQ 42

RESULT 14
AC3464
virulence-associated protein E [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AC3464
R:DeIvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3464
A:Status: preliminary
A:Molecule type: DNA

```

```

A:Residues: 1-293 <KUR>
A:Cross-references: GB:AE008917; PIDN:AA152878.1; PID:g17983722; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11697
A:Map position: I
C:Superfamily: Sphingomonas aromaticivorans hypothetical protein 229

Query Match          52.1%; Score 38; DB 2; Length 293;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
||||| |||
Db 70 EGRYTPSPPE 79

RESULT 15
T02602
vacuolar sorting receptor protein homolog At2g14740 - Arabidopsis thaliana
M:Alternate names: vacuolar sorting receptor protein homolog F26C24.12
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02602; H84520
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A:Reference number: Z14680
A:Accession: T02602
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-628 <ROU>
A:Cross-references: EMBL:AC004705; NID:g3252804; PID:g3252813
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Koflat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: H84520
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-628 <STO>
A:Cross-references: GB:AE002093; NID:g3252813; PIDN:AMC24183.1; GSPDB:GN00139
C:Genetics:
A:Gene: F26C24.12; At2g14740
A:Map position: 2
A:Introns: 102/1; 333/1; 362/3; 386/1; 414/1; 439/3; 467/1; 494/3; 517/1; 560/1; 596/

Query Match          52.1%; Score 38; DB 2; Length 628;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 EGRYTAPOPO 14
||||| |||
Db 258 KGRYCAPDPE 267

```

Search completed: January 13, 2003, 09:57:11
Job time : 14.9231 secs

Mon Jan 13 10:35:12 2003

us-09-554-941-11.rsp

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 53.1282 Seconds
(without alignments)
54.296 Million cell updates/sec

Title: US-09-554-941-11

Perfect score: 73
Sequence: 1 OASIEGRTAPOPQ 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL.21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_ornithine:*
9: sp_phage:*
10: sp_plant:*
11: sp_rhizobium:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	45	61.6	326 10 Q9ZTWO	Q9ZTWO hordenum vul
2	43	58.9	283 5 Q9VZRO	Q9VZRO drosophila
3	43	58.9	323 5 Q9VZRO	Q9VZRO drosophila
4	43	58.9	323 5 Q9VZRO	Q9VZRO drosophila
5	42	57.5	424 16 Q9VZRO	Q9VZRO drosophila
6	42	57.5	424 16 Q9VZRO	Q9VZRO drosophila
7	42	57.5	424 16 Q9VZRO	Q9VZRO drosophila
8	41	56.2	1598 5 Q9VZRO	Q9VZRO drosophila
9	40	54.8	302 5 Q9VZRO	Q9VZRO drosophila
10	40	54.8	302 5 Q9VZRO	Q9VZRO drosophila
11	40	54.8	302 5 Q9VZRO	Q9VZRO drosophila
12	40	54.8	302 5 Q9VZRO	Q9VZRO drosophila
13	39	53.4	1263 5 Q9VZRO	Q9VZRO drosophila
14	39	53.4	1263 5 Q9VZRO	Q9VZRO drosophila
15	39	53.4	1263 5 Q9VZRO	Q9VZRO drosophila
16	39	53.4	1263 5 Q9VZRO	Q9VZRO drosophila

17	39	53.4	251 13 Q8UPP3	Q8UPP3 xenopus lae
18	39	53.4	308 16 Q9L008	Q9L008 streptomyces
19	39	53.4	353 13 Q9W702	Q9W702 xenopus lae
20	39	53.4	401 5 Q8SX39	Q8SX39 streptomyces
21	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
22	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
23	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
24	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
25	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
26	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
27	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
28	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
29	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
30	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
31	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
32	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
33	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
34	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
35	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
36	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
37	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
38	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
39	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
40	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
41	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
42	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
43	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
44	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien
45	39	53.4	507 4 Q9VZRO	Q9VZRO homo sapien

ALIGNMENTS

RESULT 1
Q9ZTWO PRELIMINARY: PRT: 326 AA.

Q9ZTWO: 01-MAY-1999 (TREMREL. 10, Created)
AC 01-MAY-1999 (TREMREL. 10, Last sequence update)
DT 01-MAY-1999 (TREMREL. 21, Last annotation update)
DE ABA-responsive protein.
OS Hordenum vulgare (Barley).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poaceae:
OC Triflorae: Hordenum.
OC NCBI:Taxid=4513;
RN [1]
RP Liu J.-H., Luo M., Cheng K.-J., Mohapatra S.S., Hill R.D.:
RA "Identification and characterization of a novel barley gene that is
RT ABA-inducible and expressed specifically in embryo and aleurone."
RI J. Exp. Bot. 50:727-728(1999).
DR EMBL: AF026538; AD09343.1; GRAM.dom.
DR InterPro: IPR004182; GRAM.dom.
DR Pfam: PF02893; GRAM.1
SQ SEQUENCE 326 AA: 34301 MW: 5A6653196BD6AADE CRC64:

Query Match 61.6% Score 45; DB 10; Length 326;
Best local Similarity 70.0% Pred. No. 5.9; 1; Indels 0; Gaps 0;
Matches 7; Conservative

QY 4 IEGRTAPOP 13
DB 76 IEGRTAPOP 85
RESULT 2
Q9VZRO PRELIMINARY: PRT: 283 AA.
ID Q9VZRO: 01-MAY-2000 (TREMREL. 13, Created)
AC 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

01-JUN-2000 (TREMBlrel. 14, Last annotation update)
CG12016.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
NCBI_Taxid=7227;
SEQUENCE FROM N.A.
STRAIN-BERKELEY;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Wan K.H., Doyle C., Rogers Y.-H.C., Blazer V., Chai M., Pfeiffer B.D.,
Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Balaban D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
Borokov A., Botchan M.R., Bouck B., Brinkman P., Brotter P.,
Burris K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Paolo R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dudbin K.J., Evangelista C.C., Ferraz C., Fertara S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glocke A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ikegami C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei L., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Matthei B., McIntosh T.C., McLeod M.P., McPherson D.,
Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Moshrefi A.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
Slytskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan K.C., Wu D., Yang S., Yao Q.A.,
Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng L.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.,
Science 287:2185-2195(2000).
EMBL: AE003477; AF47759.1;
FlyBase: FBgn0035436; CG12016.
SEQUENCE 283 AA; 33018 MW; FF397060D499BC90 CRC64;
Query Match
Best Local Similarity 58.9%; Score 43; DB 5; Length 283;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYTAPQPO 14
DB 93 IKGRIHVAPEPO 103
RESULT 3
ID 095SV2 PRELIMINARY; PRT: 323 AA.
AC 095SV2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE SD05789P.
GN CG12016.
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophillidae; Drosophila.
NCBI_Taxid=7227;
SEQUENCE FROM N.A.
STRAIN-BERKELEY;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Wan K.H., Doyle C., Rogers Y.-H.C., Blazer V., Chai M., Pfeiffer B.D.,
Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Balaban D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
Borokov A., Botchan M.R., Bouck B., Brinkman P., Brotter P.,
Burris K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Paolo R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dudbin K.J., Evangelista C.C., Ferraz C., Fertara S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glocke A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Mei M.-H., Ikegami C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei L., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Matthei B., McIntosh T.C., McLeod M.P., McPherson D.,
Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Moshrefi A.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
Slytskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan K.C., Wu D., Yang S., Yao Q.A.,
Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng L.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Smith H.O.,
Science 287:2185-2195(2000).
EMBL: AE003477; AF47759.1;
FlyBase: FBgn0035436; CG12016.
SEQUENCE 283 AA; 33018 MW; FF397060D499BC90 CRC64;
Query Match
Best Local Similarity 58.9%; Score 43; DB 5; Length 283;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 IEGRYTAPQPO 14
DB 93 IKGRIHVAPEPO 103
RESULT 4
ID 09KZD6 PRELIMINARY; PRT: 525 AA.
AC 09KZD6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE Hypothetical protein SC06953.
GN SC06953 OR SC67.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
NCBI_Taxid=1902;
SEQUENCE FROM N.A.
STRAIN-A3(2)/M45;
Bentley S.D., Chater K.F., Cerdano-Parraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Kleser H.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Collins M.,
Huang C.H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
Hopwood D.A.,
"Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)."
Nature 417:141-147(2002).
EMBL: AL353870; CAB85014.1;
Hypothetical protein.
SEQUENCE 525 AA; 57338 MW; 4A7BE97C8CAEB24 CRC64;
Query Match
Best Local Similarity 72.7%; Score 43; DB 16; Length 525;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 STEGRYAPQPO 13
DB 451 SIRGIFAPQPO 461
RESULT 5
ID 059995 PRELIMINARY; PRT: 424 AA.
AC 059995;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Modification methylase (EC 2.1.1.73) (Cytosine-specific
DE methyltransferase).
GN S1R0214.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria: Cyanobacteria: Chroococcales: Synechocystis.
NCBI_TaxID=1148;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN-PCC 6803;
RA Tabata S., (Aug-1995) to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RX MEDLINE=96127523; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 648 to 928 of the genome.";
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,
RA Tabata S.;
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-
CC ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCTYOSINE.
CC -1- SIMILARITY: TO C-5 CYTOSINE-SPECIFIC DNA METHYLASE FAMILY.
CC EMBL: D64000; BAA10211.1; -;
DR HSSP: P20589; IDCT.
DR InterPro: IPR001525; C5_DNA_meth.
DR Pfam: PF00145; DNA_methylase; 1.
DR PRINTS: PR00105; C5METHYLASE.
DR TIGRfam: TIGR00675; dcm; 1.
DR PROSITE: PS00094; C5_MTHASE; 2; 1.
DR PROSITE: PS00095; C5_MTHASE; 2; 1.
KM Methyltransferase; Restriction system; Complete proteome.
SQ SEQUENCE 424 AA; 47352 MW; D0428C8C8FCC230 CRC64;
Query Match 57.5%; Score 42; DB 16; Length 424;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 2 ASIEGRYTAPOP 13
ID 1:111111
DB 319 ASDKGATYAPR 330
RESULT 6
OY99Y3 PRELIMINARY; PRT; 475 AA.
ID 0949Y3
AC Q949Y3
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative acid phosphatase.
GN A5G34850.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
OC eucosids II: Brassicales: Brassicaceae: Arabidopsids.
NCBI_TaxID=3702;
OX [1]
RN SEQUENCE FROM N.A.
RP Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene T5B15_10/At5g34850.";
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.; Clones.";
RA "Arabidopsis Open Reading Frame (ORF) Clones.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + PHOSPHATE.
CC EMBL: AY050812; AK92747.1; -;
DR EMBL: AY091415; AAM14354.1; -;
DR InterPro: IPR003961; RN_III.
DR InterPro: IPR004843; N_Peptidease.
DR InterPro: IPR004844; S/T-phosphatase.
DR Pfam: PF00149; Metallophos; 1.
DR SMART: SM00060; FN3; 1.
KM Hydrolase; Iron; Zinc.
SQ SEQUENCE 475 AA; 55009 MW; 7C73A161EE16327E CRC64;
Query Match 57.5%; Score 42; DB 10; Length 475;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 QASIEGRYTAPOP 13
ID 1:111111
DB 391 QEGLAGRTPEPOP 403
RESULT 7
OY99D99 PRELIMINARY; PRT; 590 AA.
ID 09LD99
AC 09LD99
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical protein.
DE Oryza sativa (Rice).
OS Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
NCBI_TaxID=4530;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. NIPPONBARE;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:PO469B05.";
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. NIPPONBARE;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:PO511C01.";
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AB002480; BAA96555.1; -;
DR EMBL: AB002070; BAA95849.1; -;
DR InterPro: IPR003662; sub_transporter.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
SQ SEQUENCE 590 AA; 63418 MW; DBD9BFA0CA165E CRC64;

Query Match
 Best Local Similarity 57.5%; Score 42; DB 10; Length 590;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 QASIEGRTAPOP 13
 DB 535 QRETEGRTAPOP 547

RESULT 8
 ID 09R088 PRELIMINARY; PRT; 224 AA.
 AC 09R088;
 DT 01-MAY-2000 (TREMblrel, 13, Created)
 DT 01-MAY-2000 (TREMblrel, 13, Last sequence update)
 DE 01-JUN-2001 (TREMblrel, 17, Last annotation update)
 GN MAPH dehydrogenase I, C subunit.
 OS DR1504.
 OC Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 RN [1]
 RX NCBI_TaxID=1293;
 RC SEQUENCE FROM N.A.
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moritz K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utechtack T., Zaleski C.,
 RA Kechum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1."
 RL Science 286:1571-1577(1999).
 DR EMBL: AEO01994; AAF11071.1;
 DR TRIP: DR1504;
 DR InterPro: IPR001268; Complex1_30K;
 DR Pfam: PF00329; Complex1_30Kd_1;
 DR ProDom: PD001581; Complex1_30K; 1;
 DR PROSITE: PS00542; Complex1_30K; 1;
 KW Complete proteome.
 SO SEQUENCE 224 AA; 25032 MW; 05CC3BDF11281B35 CRC64;
 Query Match
 Best Local Similarity 56.2%; Score 41; DB 16; Length 224;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 7 RYTAPOPO 14
 DB 80 RYTAPOPO 87

RESULT 9
 ID 095YM8 PRELIMINARY; PRT; 1598 AA.
 AC 095YM8;
 DT 01-DEC-2001 (TREMblrel, 19, Created)
 DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
 DE MblK-1 protein.
 GN MBLK-1.
 OS Apis mellifera (Honeybee).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OC Apoidea; Apidae; Apis.
 GN NCBI_TaxID=7460;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Takeuchi H., Kage E., Sawata M., Kamikouchi A., Ohashi K., Ohara M.,
 RA Fujiyuki T., Kunieda T., Sekimizu K., Natori S., Kubo T.;

RT "Identification of a novel gene, MblK-1, that encodes a putative
 RT transcription factor expressed preferentially in the large-type Kenyon
 RT cells of the honey bee brain."
 RL Insect Mol. Biol. 10:487-494(2001).
 DR EMBL: AB047034; BAB64310.1;
 SO SEQUENCE 1598 AA; 174929 MW; E5475BDD3ACB1EEF CRC64;

Query Match
 Best Local Similarity 56.2%; Score 41; DB 5; Length 1598;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 6 GRTAPOP 13
 DB 995 GRTAPOP 1002

RESULT 10
 ID 095RE3 PRELIMINARY; PRT; 302 AA.
 AC 095RE3;
 DT 01-DEC-2001 (TREMblrel, 19, Created)
 DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
 DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
 GN CG10263.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez N., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Paclet J., Paragas V., Park S., Phouenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL: AY061441; AAL28989.1;
 DR Flybase: FBgn0032812; CG10263.
 DR InterPro: IPR001841; Znf_ring.
 DR Pfam: PF00097; Zf-CRHC4; 1;
 DR PROSITE: PS00518; ZF_RING_1; UNKNOWN_1.
 KW Zinc-finger.
 SO SEQUENCE 302 AA; 32559 MW; 7A54C93A1FF32BC CRC64;

Query Match
 Best Local Similarity 54.8%; Score 40; DB 5; Length 302;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 1 QASIEGRTAPOP 13
 DB 248 QAHINHRVAPOP 260

RESULT 11
 ID 09VIT1 PRELIMINARY; PRT; 452 AA.
 AC 09VIT1;
 DT 01-MAY-2000 (TREMblrel, 13, Created)
 DT 01-MAY-2000 (TREMblrel, 13, Last sequence update)
 DE CG10263 protein.
 GN CG10263.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Artil J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Bottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbalt W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jajall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AF003663; AAF53834.1; -
DR Flybase: FBgn0032812; CG10263.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; ZF-C3HC4; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
KW Zinc-finger.
SQ SEQUENCE 452 AA; 47001 MW; 9763885B668FA69 CRC64;

Query Match 54.8%; Score 40; DB 5; Length 452;
Best Local Similarity 61.5%; Pred. No. 70;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 QASIEGRYAPQP 13
DB 248 QAHINRHVAFDP 260

RESULT 12
Q9A0E7 PRELIMINARY: PRT; 736 AA.
AC Q9A0E7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein SPY0807.
GN SPY0807.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBL_TaxID=1314;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RA MEDLINE-21192684; PubMed-11296296;
RA Ferretti J.J., Koshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seze G., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006531; AK33744.1; -
DR InterPro: IPR00917; Sulfatase.
DR Pfam: PF00884; Sulfatase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 736 AA; 82914 MW; A3228642886858EA CRC64;

Query Match 54.8%; Score 40; DB 16; Length 736;
Best Local Similarity 63.6%; Pred. No. 12e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 IEGRYAPQP 14
DB 1154 LEPRYSAPTPQ 1164

RESULT 14
Q29049 PRELIMINARY: PRT; 135 AA.
AC Q29049;
DT 02-NOV-2001 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Heterocyst glycolipid synthase.
GN HGLS OR AL11646.
OS Anabaena sp. (Strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBL_TaxID=103690;
RN [1]
SQ SEQUENCE FROM N.A.
RP MEDLINE-21595285; PubMed-11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Kishida Y., Iritaguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Nakazaki N., Shimp S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium *Anabaena* sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003586; BAB78012.1; -
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR000794; ketoacyl-synt.
DR Pfam: PF00698; Acyl_transf. 1.
DR Pfam: PF00109; ketoacyl-synt. C; 1.
DR Pfam: PF02801; ketoacyl-synt. C; 1.
DR Pfam: PF00350; pp-binding; 1.
DR Pfam: PF00507; ACP_DOMAIN; 1.
DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1263 AA; 137716 MW; FBA1BFA4A08B1430 CRC64;

Query Match 54.8%; Score 40; DB 16; Length 1263;
Best Local Similarity 63.6%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IEGRYAPQP 14
DB 1154 LEPRYSAPTPQ 1164

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKLETAL MUSCLE;
RX MEDLINE=99307156; PubMed=10375634;
RA Davoli R., Zambonielli P., Bigli D., Fontanesi L., Russo V.;
RT "Analysis of expressed sequence tags of porcine skeletal muscle.";
RL Gene 233:181-188(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKLETAL MUSCLE;
RA Davoli R., Fontanesi L., Cepica S., Musilova P., Stratil A.,
RA Rubens J.;
RT "The porcine poly(rC)-binding protein 2 (PCBP2) gene maps to
RT chromosome 5.";
RL Annu. Genet. 30:161-168(1999).
DR EMBL; X94253; CNA63936.1; -.
DR HSSP; Q07244; 1KHM.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH-domain; 1.
DR SMART; SM00322; KH; 1.
DR PROSITE; PS50084; KH_type_1; 1.
KM Nucleocapsid; Ribonucleoprotein.
FT NON_TER
SQ SEQUENCE 135 AA; 14144 MW; 8789F8CF3D23F2FC CRC64;

Query Match 53.4%; Score 39; DB 6; Length 135;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SIEGRYTAPOP 13
Db 14 TIQGVATIRPOP 24

RESULT 15
Q9VPZ0
ID 09VPZ0 PRELIMINARY; PRT: 179 AA.
AC 09VPZ0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG5139 protein.
GN CG5139
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews Pankoch C., Baldwin D.,
RA Balieu R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

```

```

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz L., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003587; AAF51394.1; -.
DR FlyBase; FBgn0031323; CG5139.
SQ SEQUENCE 179 AA; 21086 MW; E842E509CC016E8 CRC64;

Query Match 53.4%; Score 39; DB 5; Length 179;
Best Local Similarity 53.8%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QASIEGRYTAPOP 13
Db 51 EAETIGATIRPNP 63

```

Search completed: January 13, 2003, 09:59:48
Job time : 56.1282 secs

OM prote-
Run on:
US-06-554-0.5
422922 residues

US-0422- parameters:

262574

Title: score: 1
Perfect:
Sequence:
table: 2000000000

Score	Match	Summary
1008	Match	Summary

Search Listing
AA: *
/5A-COMB.pep: *
dep: *

```

1: /cgn2_6/ptcodata/1/1aa/9A_COMB.pep:*
2: /cgn2_6/ptcodata/1/1aa/6B_COMB.pep:*
3: /cgn2_6/ptcodata/1/1aa/PTOTUS_COMB.pep:*
4: /cgn2_6/ptcodata/1/1aa/backfile1.pep
5: /cgn2_6/ptcodata/1/1aa/backfile1.pep
6: /cgn2_6/ptcodata/1/1aa/backfile1.pep

```

Pred. No. 12 than of 13 of the
score greater by analysis of
and is derived by analysis of
SUMMARIES

Description	Quantity	Unit Price	Total Price
1. 1000 units of Product A	1000	\$10.00	\$10,000.00
2. 500 units of Product B	500	\$20.00	\$10,000.00
3. 200 units of Product C	200	\$50.00	\$10,000.00
Total: 1700 units, \$30,000.00			

Result No.	Score	Match	Length	DB	ID
1	45	61	321	4	US-09-417-286-4
2	46	49	229	1	US-08-028-463-2
3	36	49	330	1	US-08-414-391-2
4	35	47	70	4	US-09-145-044-6
5	35	47	70	4	US-09-133-887-7
6	35	47	120	6	US-09-200-867-7
7	35	47	476	2	55158255-713-8
8	35	47	2647	2	US-08-593-562B-8
9	35	47	2647	2	US-08-593-113-8
10	34	46	323	1	US-08-701-989-2
11	34	46	1298	4	US-08-591-999-5
12	34	46	1298	4	US-08-590-473-2
13	34	46	1298	4	US-08-629-821A-2
14	34	46	2616	6	US-08-683-659-2
15	34	46	1555	4	52061633-809-20
16	33	45	1555	4	US-09-369-102-0
17	33	45	1555	4	US-09-667-963-20
18	33	45	1555	4	US-09-668-480-20
19	33	45	1555	4	US-09-668-486-20
20	33	45	1555	4	US-09-568-472-20
21	33	45	1555	4	US-09-568-879-20
22	33	45	1555	4	US-09-567-889-20
23	33	45	273	3	US-09-703-191A-3
24	33	45	273	3	US-08-753-476-68
25	33	45	302	3	US-08-037-190-47
26	33	45	302	3	US-09-037-192-4
27	33	45	302	3	

```

description
sequence 4, Appl1
sequence 2, Appl1
sequence 2, Appl1
sequence 6, Appl1
sequence 7, Appl1
sequence No. 551482
Patent No. 4, Appl1
sequence 8, Appl1
sequence 8, Appl1
sequence 2, Appl1
sequence 5, Appl1
sequence 2, Appl1
sequence 2, Appl1
sequence 2, Appl1
Patent No. 20, Appl1
sequence 20, Appl1
sequence 20, Appl1
sequence 20, Appl1
sequence 20, Appl1
sequence 20, Appl1
sequence 32, Appl1
sequence 32, Appl1
sequence 68, Appl1
sequence 47, Appl1
sequence

```

28	33	45	2	302	4	US-09-037-143-47
29	33	45	2	302	4	US-09-040-691-47
30	33	45	2	302	4	US-09-040-174-47
31	33	45	2	302	4	US-08-280-128A-47
32	33	45	2	302	4	US-09-539-346-47
33	33	45	2	302	4	US-09-537-192-47
34	33	45	2	302	4	US-09-037
35	33	45	2	302	4	US-09-033
36	33	45	2	302	4	524084933-593A-8
37	33	45	2	388	4	US-07-393-593A-8
38	33	45	2	424	1	US-07-820-011A-47
39	33	45	2	530	4	US-08-426-509A-13
40	33	45	2	530	4	US-08-933-004A-47
41	33	45	2	536	5	PCT-US95-05008-12
42	33	45	2	536	5	PCT-US95-05008-11
43	33	45	2	536	5	PCT-US95-05008-11
44	33	45	2	537	4	US-08-426-509A-11
45	33	45	2	537	4	US-08-426-509A-14
46	33	45	2	543	5	PCT-US99-05008-14
47	33	45	2	543	5	PCT-US99-05008-14

ALIGNMENTS

RESULT 1
US-09-417-286-4 Application US/09417286

Sequence 4, 630665b
 Patent No. 630665b
 GENERAL INFORMATION:
 Applicant: Liu, Jin-Hao
 Applicant: Cheng, Kuo-Joan
 Applicant: Chen, Yeh-Chin
 Applicant: Chen, Yeh-Chin
 TITLE OF INVENTION: PLANT EMERGO - AND ALEURONE-SPECIFIC PROMOTER
 FILE REFERENCE: 08919-019001
 CURRENT FILING DATE: 05/09/417,286
 CURRENT FILING DATE: 1999-10-13
 NUMBER OF SEQ ID NOS: 8
 CURRENT FILING DATE: 1999-10-13
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: FASTSEQ for Windows Version 4.0
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO: 321
 LENGTH: 321
 TYPE: PRT Hordeum vulgare
 ORGANISM: Hordeum vulgare
 US-09-417-286 4
 61.6%; Score 45; DB 4; Length 321;
 Pred. NO. 2.9; Indels
 70.0%; Mismatches

Matches	
4	LEGRTAPQP 13
QY	: :
75	LEGRTAPQP 84
Db	

[illegible]

Mon Jan 13 10:35:11 2003

us-09-554-941-11.ra1

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION: Patent Release #1.0, Version #1.25
FILING DATE: 09-MAR-1993
CLASSIFICATION NUMBER: 09-08/028,463
APPLICATION NUMBER: 435
PRIORITY NUMBER: 07/694,747
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
TELEPHONE: 212 790-9090
TELEFAX: 212 869-9741
SEQUENCE CHAR SET ID NO: 2
LENGTH: 229 amino acids
STRANDEDNESS: single
MOLECULE TYPE: linear
ORIGINAL SOURCE: peptide
ORGANISM: Rhodococcus rhodochrous
STRAIN: J-1 (FERM BP-1478)
US-08-028-463-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 1; Length 229;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 1 QASIEGRTYRPOP 13
DB 99 QETLEGRTYDRKP 111

US-09-145-391-2
Sequence 2; Application US/09145391
Patent No. 6194171
GENERAL INFORMATION:
APPLICANT: Pulst, Stefan M.
TITLE OF INVENTION: Nucleic Acids Encoding Ataxin-2 Binding Proteins,
FILE REFERENCE: CE 3063
CURRENT FILING DATE: 1998-09-01
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2 Patent In Ver. 2.0
LENGTH: 330
TYPE: PPT
ORGANISM: Homo sapiens
US-09-145-391-2

Query Match
Best Local Similarity 49.3%; Score 36; DB 4; Length 330;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 1 QASIEGRTYRPOP 13
DB 34 QNGCIPAEYTRHP 46

US-09-133-944-6
Sequence 6; Application US/09133944
Patent No. 6280937
GENERAL INFORMATION:
APPLICANT: Luo, Ying
TITLE OF INVENTION: SHUTTLER VECTORS
FILE REFERENCE: A66522/DJB/DIV
CURRENT FILING DATE: 1999-08-14
US-09-133-944-6

Mon Jan 13 10:35:11 2003

us-09-554-941-11.ra1

```

; EARLIER APPLICATION NUMBER: 09/133,949
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: minbody
; OTHER INFORMATION: presentation structure
; US-09-133-944-6

```

```

Query Match 47.9%; Score 35; DB 4; Length 70;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 ASIEGRYT 9
   11:11111
Db 43 ASYKGRYT 50

```

```

RESULT 6
US-09-208-827-7
Sequence 7, Application US/09208827

```

```

; Patent No. 6391582
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Yu, Pei Wen
; APPLICANT: Lotens, James
; TITLE OF INVENTION: SHUTTLE VECTORS
; FILE REFERENCE: A66252-1/DJB/DAV
; CURRENT APPLICATION NUMBER: US/09/208,827
; EARLIER FILING DATE: 1998-12-09
; CURRENT FILING DATE: 1998-08-14
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: minbody
; OTHER INFORMATION: presentation structure
; US-09-208-827-7

```

```

Query Match 47.9%; Score 35; DB 4; Length 70;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 ASIEGRYT 9
   11:11111
Db 43 ASYKGRYT 50

```

```

RESULT 7
551582-36
; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989

```

```

; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 36
; LENGTH: 120
; 5514582-36

```

```

Query Match 47.9%; Score 35; DB 6; Length 120;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 QASIEGRYTAPOPO 14
   11:11111
Db 44 QCTSEGVWTPVPQ 57

```

```

RESULT 8
US-08-955-713-4
Sequence 4, Application US/08955713

```

```

; Patent No. 5955308
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
; APPLICANT: MOONEY, JEFFREY
; APPLICANT: BERGSMAN, DEBK
; APPLICANT: HALSLEY, WENDY
; TITLE OF INVENTION: CDNA CLONE HEOAD54 THAT ENCODES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P. O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,713
; FILING DATE: 23-OCT-1997
; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,124
; FILING DATE: 18-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169

```

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-955-713-4

```

```

Query Match 47.9%; Score 35; DB 2; Length 476;
Best Local Similarity 55.6%; Pred. No. 2,3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 5 EGRYTAPOP 13
   11:11111
Db 356 QGQHTVPOP 364

```

```

RESULT 9
US-08-583-562B-8

```

Sequence 8, Application US/08583562B
Patent No. 5922570

GENERAL INFORMATION:

APPLICANT: Staunton, Donald
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/583,562B
FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-583-562B-8

Query Match

Best Local Similarity 47.9%; Score 35; DB 2; Length 2647;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 YTAPOP 13
|||||

Db 1712 YTAPOP 1717

RESULT 10

US-08-779-113-8
Sequence 8, Application US/08779113
Patent No. 5948891

GENERAL INFORMATION:

APPLICANT: Staunton, Donald E.
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,113
FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 5948891and
REGISTRATION NUMBER: 35,302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: not relevant
US-08-779-113-8

Query Match

Best Local Similarity 47.9%; Score 35; DB 2; Length 2647;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 YTAPOP 13
|||||

Db 1712 YTAPOP 1717

RESULT 11

US-08-591-989-2
Sequence 2, Application US/08591989
Patent No. 5795721

GENERAL INFORMATION:

APPLICANT: Ross S. Rabin, Sumedha Jayasena
TITLE OF INVENTION: and Larry Gold
TITLE OF INVENTION: HIGH AFFINITY NUCLEIC
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESS: Swanson & Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,989
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3433
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 235
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-591-989-2

Query Match

Best Local Similarity 46.6%; Score 34; DB 1; Length 235;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRAYPOP 14
|||||

Db 26 GRTAGOPR 34

RESULT 12

US-08-591-989-5

; Sequence 5, Application US/08591989

; Patent No. 5795721

; GENERAL INFORMATION:

; APPLICANT: Ross S. Rabin, Smedha Jayasena

; APPLICANT: and Larry Gold

; TITLE OF INVENTION: HIGH AFFINITY NUCLEIC

; TITLE OF INVENTION: ACID LIGANDS OF ICP4

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Swanson & Bratschun, L.L.C.

; STREET: 8400 East Prentice Avenue, Suite #200

; CITY: Englewood

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.40 MB

; MEDIUM TYPE: storage

; COMPUTER: IBM COMPATIBLE

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/591,989

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Barry J. Swanson

; REGISTRATION NUMBER: 33,215

; REFERENCE/DOCKET NUMBER: NEX 49

; TELECOMMUNICATION INFORMATION:

; TELEFAX: (303) 793-3433

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 323

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-591-989-5

; Query Match

; Best Local Similarity 46.6%; Score 34; DB 1; Length 323;

; Best Local Similarity 66.7%; Pred. No. 2.2e+02;

; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAGOPR 14

; Sequence 2, Application US/08690473

; Patent No. 5876923

; GENERAL INFORMATION:

; APPLICANT: Leopardi, Rosario

; APPLICANT: Roizman, Bernard

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN

; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690,473

FILING DATE: 26-JUL-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: ARCD:239

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/474-7577

TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1298 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-690-473-2

Query Match

; Best Local Similarity 46.6%; Score 34; DB 2; Length 1298;

; Best Local Similarity 66.7%; Pred. No. 1e+03;

; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAGOPR 14

; Sequence 2, Application US/09259821A

; Patent No. 6210926

; GENERAL INFORMATION:

; APPLICANT: LEOPARDI, ROSARIO

; APPLICANT: ROIZMAN, BERNARD

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS

; FILE REFERENCE: ARCD:317

; CURRENT APPLICATION NUMBER: US/09/259,821A

; CURRENT FILING DATE: 1999-03-01

; PRIOR APPLICATION NUMBER: 08/690,473

; PRIOR FILING DATE: 1996-07-26

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1298

; TYPE: PRT

; ORGANISM: HERPES VIRUS, TYPE 1

; US-09-259-821A-2

Query Match

; Best Local Similarity 46.6%; Score 34; DB 4; Length 1298;

; Best Local Similarity 66.7%; Pred. No. 1e+03;

; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 GRTAGOPR 14

; Sequence 2, Application US/08843659

; Patent No. 6218103

; GENERAL INFORMATION:

; APPLICANT: Leopardi, Rosario

; APPLICANT: Roizman, Bernard

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS

; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,659
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:519
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-843-659-2

Query Match 46.6%; Score 34; DB 4; Length 1298;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 GRYTAPQPQ 14
||:|:|:
Db 281 GRFTAGQPR 289

Search completed: January 13, 2003, 09:56:26
Job time: 14.2051 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:15 ; Search time 14 Seconds
(without alignments)
41.476 Million cell updates/sec

Title: US-09-554-941-11
Perfect score: 73
Sequence: 1 QASISGRYTAPOPO 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	57.5	218	CLP2_STRCO	092h58 streptomyc
2	39	53.4	362	PCB2_MOUSE	061990 mus musculu
3	39	53.4	365	PCB2_HUMAN	015366 homo sapien
4	39	53.4	423	WDRC_HUMAN	099217 homo sapien
5	39	53.4	423	WDRC_MOUSE	091144 mus musculu
6	39	53.4	508	FAS3_DROME	P15276 drosophila
7	38	52.1	663	TRA_BPMU	P07636 bacterioph
8	37	50.7	138	DIU1_MANSE	P21819 manduca sex
9	37	50.7	768	LEM3_MOUSE	001102 mus musculu
10	37	50.7	768	LEM3_RAT	P98106 rattus norv
11	37	50.7	769	LEM3_SHEEP	P98109 ovvis aries
12	37	50.7	798	HMPB_DROME	P31264 drosophila
13	37	50.7	798	YP73_CABEL	009222 caenorhabdi
14	36	49.3	224	PURQ_MYCTU	P71641 mycobacteri
15	36	49.3	229	NHBI_HOHRH	P21220 rhodococcus
16	36	49.3	363	VRAB_ECOLI	P43319 escherichia
17	36	49.3	396	A2BP_MOUSE	091143 mus musculu
18	36	49.3	397	A2BP_HUMAN	009255 caenorhabdi
19	36	49.3	469	YOB1_CABEL	009225 caenorhabdi
20	36	49.3	484	OAR1_LOCOMI	025321 locusta mig
21	36	49.3	484	OAR2_LOCOMI	025322 locusta mig
22	36	49.3	531	SRCL_XENLA	P13115 xenopus lae
23	36	49.3	531	SRCL_XENLA	P13116 xenopus lae
24	36	49.3	536	FNX_XIPHE	P27446 xiphophorus
25	36	49.3	543	PKRA_STRCO	P27446 xiphophorus
26	36	49.3	543	PKRA_STRCO	P27446 xiphophorus
27	35.5	48.6	221	GRPE_BACST	O59240 bacillus st
28	35.5	48.6	680	GAOA_DACDE	O01745 dactylium d
29	35.5	48.6	1695	KFLA_MOUSE	P13117 mus musculu
30	35	47.9	241	PCXB_ACICA	P20372 actinobact
31	35	47.9	271	YLAB_ECOLI	P77518 escherichia
32	35	47.9	324	KIME_AERPE	O95946 aeropyrum p
33	35	47.9	331	DIU1_ECOLI	P15286 escherichia

Result	ID	CLP2_STRCO	STANDARD	PRT	218 AA	ALIGNMENTS
34	35	47.9	333	1	CAP1_DICDI	P19198 dictyosteli
35	35	47.9	425	1	L756_CABEL	O1184 caenorhabdi
36	35	47.9	682	1	RHSE_ECOLI	P24211 escherichia
37	35	47.9	802	1	XYND_RUMFL	O53317 ruminoococc
38	35	47.9	811	1	Y104_SYNY3	P54371 synneococyt
39	35	47.9	902	1	GPC2_HUMAN	O58512 homo sapien
40	35	47.9	1033	1	CR2_HUMAN	P20093 homo sapien
41	35	47.9	1531	1	NP75_HUMAN	O94916 homo sapien
42	35	47.9	1647	1	PLNA_HUMAN	P21333 homo sapien
43	34.5	47.3	651	1	ABG5_HUMAN	O9h222 homo sapien
44	34.5	47.3	1690	1	KFLA_HUMAN	O12756 homo sapien
45	34	46.6	154	1	CYCP_CHRVI	P00154 chromatium

RESULT 1
CLP2_STRCO
ID CLP2_STRCO STANDARD PRT 218 AA
AC 092h58; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease proteolytic subunit 2 (EC 3.4.21.92)
DE (Phosphatase Clp 2).
DE CLP2 OR SC02618 OR SCC80.03C.
GN Streptomyces coelicolor.
OS Bacteria; Actinobacteria; Actinobacteridae; Streptomyces.
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=99255687; PubMed=10320574;
RA de Crey-Lagard V., Servant-Molsson P., Viala J., Grandvalet C., Mazodier P.;
RT *Alteration of the synthesis of the Clp ATP-dependent protease affects morphological and physiological differentiation in Streptomyces";
RL Mol. Microbiol. 32:505-517(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Bontley S.D., James K.D., Harris D.E., Quail M.A., Kieser H., Thomson N.R., Brown S., Chandra G., Chen C.W., Collins M., Harper D., Bateman A., Goble A., Hidalgo J., Hornsby T., Howarth S., Cronin A., Fraser A., Goble A., Hildegarde J., Hornsby T., O'Neill S., Huang C.-H., Krieser T., Lathe J., Murphy L., Oliver K., O'Neill S., Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)".
RL Nature 417:141-147(2002).
-1- FUNCTION: Cleaves peptides in various proteins in a process that requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a major role in the degradation of misfolded proteins (By similarity).
-1- CATALYTIC ACTIVITY: Hydrolysis of proteins to small peptides in the presence of ATP and magnesium. Alpha-casein is the usual test substrate. In the absence of ATP, only oligopeptides shorter than five residues are cleaved (such as succinyl-Leu-Tyr-L-NHMe); and Leu-Tyr-Leu-L-Tyr-Tip, in which the cleavage of the -Tyr-L-Leu-L-Tyr-L-Tip bond also occurs).
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF071885; AAC70948.1; -
 DR EMBL: AF442143; CAC09994.1; -
 DR HSP: P19245; 1TYF.
 DR MEROPS: S14.UPW; -
 DR Interpro: IPR001907; CLP_protease.
 DR Pfam: PF00574; CLP_protease; 1.
 DR PRINTS: PR00127; CLP_protease; 1.
 DR TIGRfam: TIGR00493; CLP; 1.
 DR PROSITE: PS00382; CLP_PROTEASE_HIS; 1.
 DR PROSITE: PS00381; CLP_PROTEASE_SER; 1.
 DR Hydrolase; serine protease; Multigene family; Complete proteome.
 FT ACT_SITE 114 114
 FT ACT_SITE 139 139 BY SIMILARITY.
 SQ SEQUENCE 218 AA; 24180 MW; 67082B89EAC49F2E CRC64;

Query Match
 Best Local Similarity 57.5%; Score 42; DB 1; Length 218;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 QASIEGRYAPQP 14
 Db 2 RAASGRYTPQAE 15

RESULT 2

PCB2_MOUSE
 ID PCB2_MOUSE STANDARD; PRT; 362 AA.
 AC 061950; 061883; 062042;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 15-JUN-2002 (Rel. 35; Last sequence update)
 DE Poly(ribo)-binding protein 2 (Alpha-CP2) (Putative heterogeneous nuclear
 DE ribonucleoprotein X) (hnRNP X) (CTBP) (CBP).
 GN PCB2 OR HNRNPX OR HNRNPX OR CBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RA MEDLINE=93376518; PubMed=8367306;
 RA Hahn K.B., Kim G., Turch C., Smale S.T.;
 RT "Isolation of a murine gene encoding a nucleic acid-binding protein
 RT with homology to hnRNP K.";
 RL Nucleic Acids Res. 21:3894-3894 (1993).
 RN [2]
 RP SEQUENCE FROM N.A. AND PARTIAL SROUCE (ISOFORM 2).
 RA STRAIN=C57BL/6 X 129/Ola; TISSUE=Liver;
 RA MEDLINE=94268912; PubMed=8208614;
 RA Goller M., Funke B., Gehe-Becker C., Kroege B., Lottspeich F.,
 RA Horak I.;
 RT "Murine protein which binds preferentially to oligo-C-rich single-
 RT stranded nucleic acids.";
 RL Nucleic Acids Res. 22:1885-1889 (1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RA Horak I.;
 RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Makeyev A.V., Liehaber S.A.;
 RT "Identification of two novel mammalian genes establishes a subfamily
 RT of KH-domain RNA-binding proteins.";
 CC submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN THAT
 CC BINDS PREFERENTIALLY TO OLIGO DC.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC use by non-profit institutions. There are no restrictions on its
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send an email to license@isb-sib.ch).

DR EMBL: L19661; AAA03705.1; -
 DR EMBL: X75947; CAA53346.1; -
 DR EMBL: X79982; CAA66619.1; -
 DR EMBL: AF236845; AAK14059.1; -
 DR EMBL: AF236842; AAK14059.1; JOINED.
 DR EMBL: AF236843; AAK14059.1; JOINED.
 DR EMBL: AF236844; AAK14059.1; JOINED.
 DR HSP: 007244; 1KHM.
 DR MGD: MGI:108202; Pcbp2.
 DR Interpro: IPR004087; KH_dom.
 DR Interpro: IPR004088; KH_type-1.
 DR Pfam: PF00013; KH-domain; 3.
 DR SMART: SM00322; KH; 3.
 DR PROSITE: PS50084; KH_type_1; 3.
 DR Repeat; Alternative splicing; Ribonucleoprotein; DNA-binding;
 FT DOMAIN 13 75 KH 1.
 FT DOMAIN 97 162 KH 2.
 FT DOMAIN 194 224 KH 3.
 FT VASPLIC 194 224 MISSING (IN ISOFORM 2).
 FT VASPLIC 263 275 MISSING (IN ISOFORM 3).
 SQ SEQUENCE 362 AA; 38221 MW; 70C8AF710E3F3C0 CRC64;

Query Match
 Best Local Similarity 53.4%; Score 39; DB 1; Length 362;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Oy 3 SIGGRYAPQP 13
 Db 227 TIQGYAIPQP 237

RESULT 3

PCB2_HUMAN
 ID PCB2_HUMAN STANDARD; PRT; 365 AA.
 AC 015366;
 DT 30-MAY-2000 (Rel. 39; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DE Poly(ribo)-binding protein 2 (Alpha-CP2) (hnRNP-E2).
 GN PCB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95331278; PubMed=7607214;
 RA Leffers H., Dejgaard K., Celis J.E.;
 RT "Characterisation of two major cellular poly(ribo)-binding human
 RT proteins, each containing three K-homologous (KH) domains.";
 RL Eur. J. Biochem. 230:447-453 (1995).
 CC -1- FUNCTION: MAJOR CELLULAR POLY(RC)-BINDING PROTEIN. BINDS ALSO
 CC POLY(RU).
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND IN THE NUCLEUS. MAY SHUTTLE
 CC BETWEEN THE NUCLEUS AND THE CYTOSOL.
 CC -1- TISSUE SPECIFICITY: DETECTED IN ALL TISSUES EXAMINED.
 CC -1- PTM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE
 CC STRONGEST POLY(RC)-BINDING ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 3 KH DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X78136; CA55015.1; -
 CC HSSP; Q07244; 1KH; PCBP2.
 CC Genew; HGNC:8648; PCBP2.
 CC MIM; 601210; -
 CC InterPro: IPR004087; KH_dom.
 CC InterPro: IPR004088; KH_type_1.
 CC Pfam; PF00013; KH-domain; 3.
 CC SMART; SM00322; KH; 3.
 CC PROSITE; PS50084; RN_type_1; 3.
 CC PROSITE; PS50084; RN-binding; Ribonucleoprotein; DNA-binding;
 CC Nuclear protein; RNA-binding; Repeat.
 CC Phosphorylation; Repeat.
 CC DOMAIN 13 75 KH 1.
 CC DOMAIN 97 162 KH 2.
 CC DOMAIN 287 351 KH 3.
 CC SEQUENCE 365 AA; 38580 MW; 43F035D76FDC2C63 CRC64;
 CC SQ

Query Match 53.4%; Score 39; DB 1; Length 365;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SIEGRTAPOP 13
 :|:|:|:|
 Db 231 TIOGQYAIPOP 241

RESULT 4
 WDRG_HUMAN STANDARD: PRT; 423 AA.
 ID WDRG_HUMAN; Q96ZL7; Q9N718; Q96H00; Q9N780;
 AC Q96ZL7; Q9N718; Q96H00; Q9N780;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WD-repeat protein 12 (YTM1 homolog).
 GN WDR12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matsumoto S.;
 RA "Human homolog of Saccharomyces cerevisiae YTM1";
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niimura K., Iwayanagi T.;
 RA "NDO human cDNA sequencing project";
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Skin;
 RA Strausberg R.;
 RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC -i SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAIN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF242546; AAF60355.1; -
 CC EMBL; AK001743; BAA91875.1; -
 CC EMBL; AK022781; BAB14242.1; -
 CC EMBL; AK022782; BAB14243.1; -
 CC EMBL; BC008082; AAO8082.1; -
 CC Genew; HGNC:14098; WDR12.
 CC InterPro: IPR001680; WD40.
 CC Pfam; PF00400; WD40; 7.
 CC PRINTS; PR00320; GPROTEINRPT.
 CC PRODOM; PD000018; WD40; 7.
 CC SMART; SM00320; WD40; 7.
 CC PROSITE; PS00678; WD_REPEATS_1; 2.
 CC PROSITE; PS50082; WD_REPEATS_2; 5.
 CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC Repeat; WD repeat; Polymorphism.
 CC REPEAT 99 137 WD 1.
 CC REPEAT 138 180 WD 2.
 CC REPEAT 187 226 WD 3.
 CC REPEAT 255 293 WD 4.
 CC REPEAT 295 334 WD 5.
 CC REPEAT 340 380 WD 6.
 CC REPEAT 384 422 WD 7.
 CC REPEAT 72 72 WD 7.
 CC VARIANT 89 89 /FTID=VAR_012863.
 CC VARIANT 89 89 /FTID=VAR_012864.
 CC VARIANT 286 286 /FTID=VAR_012865.
 CC CONFLICT 75 75 I -> V (IN REF. 2; BAA91875).
 CC CONFLICT 333 333 L -> M (IN REF. 1).
 CC SEQUENCE 423 AA; 47708 MW; 6D088C640AC981D8 CRC64;
 CC SQ

Query Match 53.4%; Score 39; DB 1; Length 423;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RYTAPOPO 14
 :|:|:|:|
 Db 88 KYTAPOPE 95

RESULT 5
 WDRG_MOUSE STANDARD: PRT; 423 AA.
 ID WDRG_MOUSE; Q9JUF5; Q9CST3;
 AC Q9JUF5; Q9CST3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WD-repeat protein 12 (YTM1 homolog).
 GN WDR12.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Thymus;
 RA MEDLINE=2168157; PubMed=11827460;
 RA Nal B., Mohr E., Da Silva M.I., Tagett R., Navarro C., Carroll P.,
 RA Depetris D., Vertuhy C., Jordan B.R., Ferrer P.;
 RA "Wdr12, a mouse gene encoding a novel WD-repeat protein with a
 RA notchless-like amino-terminal domain";
 RA Genomics 79:77-86(2002).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=129/SV X C57BL/6J;
 RA Matsumoto S.;
 RA "Mouse homologue of Saccharomyces cerevisiae YTM1";
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL [3]

SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6: TISSUE-Brain;
 RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
 RA Hashimoto K.,
 RT "Isolation of full-length cDNA clones from mouse brain cDNA library
 RT made by oligo-capping method."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6T: TISSUE-Testis, and Pancreas;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasuwa T., Saito R.,
 RA Fleischmann W., Gaasterland T., Gissi C., Kling B., Kocikawa H.,
 RA Schiomi L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guelinrich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Nomberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC Strausberg R.,
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AY059431; AAL29660.1;
 DR EMBL: AY059432; AAL29661.1;
 DR EMBL: AF239765; AAF44683.1;
 DR EMBL: AB041608; BAA95091.1;
 DR EMBL: AK007500; BAB2972.1;
 DR EMBL: AK012022; BAB2979.1;
 DR EMBL: AK016611; BAB3036.1;
 DR EMBL: BC004748; AAR04748.1;
 DR MGI: 1927241; WGI12.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40.7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR PRODOM: PD000018; WD40.2.
 DR SMART: SM00320; WD40.7.
 DR PROSITE: PS00678; WD_REPEATS_1; 2.
 DR PROSITE: PS00682; WD_REPEATS_2; 5.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat, WD repeat.
 FT REPEAT 99 137 WD 1.
 FT REPEAT 138 180 WD 2.
 FT REPEAT 187 226 WD 3.
 FT REPEAT 225 293 WD 4.
 FT REPEAT 295 334 WD 5.
 FT REPEAT 340 380 WD 6.
 FT REPEAT 384 422 WD 7.
 FT REPEAT 422 499 WD 7.
 FT CONFLICT 99 99 F -> S (IN REF. 2).
 FT CONFLICT 418 418 T -> A (IN REF. 1).
 SQ SEQUENCE 423 AA; 47346 MM; DF57952320373667 CRC64;

Query Match
 Best Local Similarity 53.4%; Score 39; DB 1; Length 423;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 7 RYAPAPQ 14
 Db 88 KYTAPOPE 95
 RESULT 6
 FAST3-DROME
 ID FAST3-DROME STANDARD: PRT; 508 AA.
 AC P15278; Q9V089.
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fasciclin III precursor (FAS III).
 GN FAS3 OR CG5803
 OS Drosophila melanogaster (Fruit fly).
 OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI-TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90030406; PubMed=2509076;
 RA Snow P.M., Bieber A.J., Goodman C.S.,
 RT "Fasciclin III: a novel homophilic adhesion molecule in Drosophila."
 RL Cell 59:313-323(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkley;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Anell J.F., Adyanthi A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brinkstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., DeLong E., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Durkin K.J., Evans G.H., Garg N.S., Gelbart W.M., Glasser K.,
 RA Foster A., Gough J., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Woodegate D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Wooley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: MEDIATES CELL ADHESION IN A CA-INDEPENDENT MANNER. IT

PLAYS A ROLE IN AXON OUTGROWTH, GUIDANCE AND FASCICULATION OF THE DEVELOPING NERVOUS SYSTEM.

CC SUBCELLULAR LOCATION: Type I membrane protein.

CC TISSUE SPECIFICITY: EXPRESSED ON DIFFERENT SUBSETS OF AXON BUNDLES (FASCICLES) IN INSECT EMBRYOS.

CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation on the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: M27813; AAA28532.1; -

CC EMBL: AE003658; AAP53655.1; -

CC PIR: A33378; A33378

CC FlyBase: FBgn0000636; Fasn3.

CC Cell adhesion; Glycoprotein; Repeat; Immunoglobulin domain;

CC Transmembrane; Signal; Neurogenesis; Phosphorylation.

CC SIGNAL 1 20

CC CHAIN 21 508 FASCICLIN III.

CC DOMAIN 21 346 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 347 370 POTENTIAL.

CC CYTOPLASMIC (POTENTIAL).

CC DOMAIN 371 508 IG-LIKE V-TYPE DOMAIN.

CC DOMAIN 44 106 IG-LIKE C2-TYPE DOMAIN 1.

CC DOMAIN 143 218 IG-LIKE C2-TYPE DOMAIN 2.

CC DOMAIN 256 310 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).

CC MOD_RES 21 21 POTENTIAL.

CC DISULFID 150 211 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC MOD_RES 382 382 PHOSPHORYLATION (POTENTIAL).

CC MOD_RES 459 459 PHOSPHORYLATION (POTENTIAL).

CC CONFLICT 1 15 MSRIYFICLAAILTD -> MEKQAEISGYAMH (IN REF.

CC SEQUENCE 508 AA; 55883 MW; 6E39EA0580697DAF C6C64;

CC Query Match 53.4%; Score 39; DB 1; Length 508;

CC Best Local 50.0%; Pred. No. 18;

CC Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 QASIEGRYAPQ 14

DB 279 QGRTDGRYSAYEPQ 292

RESULT 7

TR_BPMU STANDARD: PRT: 663 AA.

AC P07636; P06021; (Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Transposase.

GN A OR 3.

OS Bacteriophage Mu.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC Mu-like viruses.

OX NCBI_Taxid=10677;

OX [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=86061968; PubMed=2999776;

RA Haishey R.M., Getzoff E.D., Baldwin D.L., Miller J.L., Chaconas G.;

RT "Primary structure of phage mu transposase: homology to mu

RT Proc. Natl. Acad. Sci. U.S.A. 82:7676-7680(1985).

RN [2]

RP SEQUENCE FROM N.A.

RA Priest H., Brauer B., Schmidt C., Kamp D.;

RT "Sequence of the left end of Mu.";

RL (in) Symonds N., Toussaint A., van de putte P., Howe M.M. (eds.);

RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,

RL New York (1987).

RL [3]

RP SEQUENCE FROM N.A.

RA Morgan G., Hatfull G., Hendrix R.;

RT "Genome of bacteriophage Mu and comparison with the haemophilus

RT Influenzae Mu-like prophage Flumu.";

RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE OF 1-88 FROM N.A.

RA MEDLINE=83012203; PubMed=6214696;

RA Priest H., Kamp D., Kahmann R., Brauer B., Deltis H.;

RT "Nucleotide sequence of the immunity region of bacteriophage Mu.";

RL Mol. Gen. Genet. 186:315-321(1982).

RL [5]

RP SEQUENCE OF 1-84 FROM N.A.

RA MEDLINE=83218562; PubMed=6222246;

RA Toussaint A., Faellen M., Desmet L., Allet B.;

RT "The products of gene A of the related phages Mu and D108 differ in

RT their specificities.";

RL Mol. Gen. Genet. 190:70-79(1983).

RL [6]

RP STRUCTURE BY NMR OF 1-76.

RA MEDLINE=95187707; PubMed=7881904;

RA Clubb R.T., Michlinski J.G., Savilahti H., Mizuuchi K.,

RA Gronenborn A.M., Clore G.M.;

RT "A novel class of winged helix-turn-helix protein: the DNA-binding

RT domain of Mu transposase.";

RL Structure 2:1041-1048(1994).

RL [7]

RP STRUCTURE BY NMR OF 76-174.

RA MEDLINE=98070329; PubMed=9405381;

RA Schumacher S., Clubb R.T., Cal M., Mizuuchi K., Clore G.M.,

RA Gronenborn A.M.;

RT "Solution structure of the Mu end DNA-binding Ibeta subdomain of

RT phage Mu transposase: modular DNA recognition by two tethered

RT domains.";

RL EMBO J. 16:7532-7541(1997).

RL [8]

RP STRUCTURE BY NMR OF 173-247.

RA MEDLINE=98035037; PubMed=9367742;

RA Clubb R.T., Schumacher S., Mizuuchi K., Gronenborn A.M., Clore G.M.;

RT "Solution structure of the I gamma subdomain of the Mu end

RT DNA-binding domain of phage Mu transposase.";

RL J. Mol. Biol. 273:19-25(1997).

RL [9]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 248-574.

RA MEDLINE=95354202; PubMed=7628012;

RA Rice P., Mizuuchi K.;

RT "Structure of the bacteriophage Mu transposase core: a common

RT structural motif for DNA transposition and retroviral integration.";

RL Cell 82:209-220(1995).

RL -1- FUNCTION: THIS TRANSPOSASE IS ESSENTIAL FOR INTEGRATION,

CC REPLICATION-TRANSPOSITION, AND EXCISION OF MU DNA.

CC -1- MISCELLANEOUS: MU CAN TRANSPOSE ITS DNA INTO MULTIPLE SITES IN

CC MANY BACTERIAL GENOMES AND MEDIANE A VARIETY OF DNA

CC REARRANGEMENTS. TRANSPOSITION REQUIRES BOTH TRANSPOSASE (ENCODED

CC BY GENE A) AND TRANSDUCTION ENHANCER (ENCODED BY GENE B).

CC -1- MISCELLANEOUS: UNLIKE OTHER TRANSPOSONS MU HAS DISTINCTLY

CC SEQUENCES AT ITS LEFT AND RIGHT ENDS. TRANSPOSASE APPARENTLY BINDS

CC 3 SPECIFIC BLOCKS OF SEQUENCES AT EACH END OF MU DNA.

CC -1- MISCELLANEOUS: THE A GENE IS REGULATED BY THE REPRESSOR C, WHICH

CC BINDS TO AN OPERATOR SEQUENCE & TURNS OFF TRANSCRIPTION. REPRESSOR

CC C CAN, AT HIGH CONCENTRATIONS, OCCUPY ALMOST THE EXACT SAME SITES

CC ON MU ENDS AS THE TRANSPOSASE, AND TRANSPOSASE CAN BIND TO

CC FRAGMENTS CONTAINING THE MU OPERATOR SEQUENCE.

CC -1- SIMILARITY: STRONG. TO H. INFLUENZAE H11478.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

DT 01-APR-1993 (Rel. 25, last sequence update)
DE 15-JUN-2002 (Rel. 41, last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADSEMP
DE (cd62p) (Leukocyte-endothelial cell adhesion molecule 3) (L3CAM3).
OS SLEP OR GRMP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92340571; PubMed=1378646;
RA Weller A., Isenmann S., Vestweber D.;
RT "Cloning of the mouse endothelial selectins. Expression of both E-
RT and P-selectin is inducible by tumor necrosis factor alpha.";
RN J. Biol. Chem. 267:15176-15183(1992).
RP SEQUENCE FROM N.A.
RX MEDLINE=92345617; PubMed=1379089;
RA Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;
RT "Molecular cloning and analysis of in vivo expression of murine p-
RT selectin.";
RL Blood 80:795-800(1992).
CC -I- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC THE CELL SURFACE.
CC -I- INDUCTION: BY TNF-ALPHA.


```

FT CARBOHYD      398      398      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     603      603      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     654      654      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     661      661      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD     679      679      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID         745      745      PALMITATE (BY SIMILARITY).
FT SITE         756      759      ENDOCYTOSIS SIGNAL (PROBABLE).
FT CONFLICT     724      724      A -> E (IN REF. 2).
SQ SEQUENCE     768 AA: 83098 MW: ES17307AD266E68 CRC64;

Query Match          50.7%; Score 37; DB 1; Length 768;
Best Local Similarity 38.5%; Pred. No. 65;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY      1 QASIEGRYAPQP 13
       : :: ||::|| |
Db      553 ECTVSGRMSAPP 565

RESULT 10
LEW3_RAT LEW3_RAT STANDARD; PRT; 768 AA.
ID LEW3_RAT
AC P98I06;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P-selectin precursor (granule membrane protein 140) (GMP-140) (PADGEN)
DE (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN SLP.
NC Rattus norvegicus (Rat.).
CC Pakayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94333817; PubMed=7520013;
RA Auchampach J.A., Oliver M.G., Anderson D.C., Manning A.M.;
RT "Cloning, sequence comparison and in vivo expression of the gene
   encoding rat P-selectin."
RL Gene 145:251-255(1994).
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
   TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
   INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
   LEUCOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
   LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.
CC -1- INDICATION: BY ACUTE INFLAMMATION (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SELECTIN/LECTIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS; RAT P-LECTIN LACKS THE
   HUMAN SUSHI-2 EQUIVALENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation on its
CC European Bioinformatics Institute. There are no restrictions on its use by
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensel@sb-sib.ch).
CC -----
CC EMBL: L23086; AAA60325.1; -
CC HSSP: P16109; IRSB
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001304; Lectin_C
CC InterPro: IPR002396; Selectin
CC InterPro: IPR000436; Susln_SCR_CCP
CC pfam: PF00059; lectin_c_1
CC pfam: PF00084; susln_8
CC PRINTS: PR00343; SELECTIN
CC SMART: SM00032; CCP_8.

```

DE	P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE	(CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN	SERP.
OS	Ovis aries (Sheep).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Caprinae; Ovis.
OX	NCBL_TaxID=940;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Heart;
RA	Burns S.A., Neufeld E.J., Donady J.J.;
RL	Submitted (JUL-1994) to the EMBL/Genbank/DDBJ databases.
CC	-I- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC	LEUKOCYTES. THE LIGAND RECOGNIZED IS STALYL-LEWIS X.
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-I- SIMILARITY: BELONGS TO THE SELECTIN/LECAM FAMILY.
CC	-I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN/FAMILY DOMAIN.
CC	-I- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: L34270; AA859261.1; ..
DR	HSSP: P6109; IESB.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR001304; Lectin.C.
DR	InterPro: IPR002396; Selectin.
DR	pfam: PF00008; EGF_1.
DR	pfam: PF00059; lectin_c; 1.
DR	Pfam: PF00084; sushi; 8.
DR	PRINTS: PR00343; SELECTIN.
DR	SMART: SM00032; CCP; 8.
DR	SMART: SM00034; CLECT; 1.
DR	SMART: PS00181; EGF; 1.
DR	PROSITE: PS00022; EGF_1; 1.
DR	PROSITE: PS01186; EGF_2; 1.
DR	PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
FT	Selectin; signal; Sushi; Repeat.
FT	STGNLT; 1 32
FT	CHAIN 1 32
FT	DOMAIN 33 769 POTENTIAL.
FT	TRANSMEM 33 717 P-SELECTIN.
FT	DOMAIN 718 734 EXTRACELLULAR (POTENTIAL).
FT	DOMAIN 735 769 POTENTIAL.
FT	DOMAIN 58 158 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 159 195 C-TYPE LECTIN.
FT	DOMAIN 199 258 EGF-LIKE.
FT	DOMAIN 261 320 SUSHI 1.
FT	DOMAIN 323 382 SUSHI 2.
FT	DOMAIN 385 444 SUSHI 3.
FT	DOMAIN 447 506 SUSHI 4.
FT	DOMAIN 509 568 SUSHI 5.
FT	DOMAIN 580 639 SUSHI 6.
FT	DOMAIN 642 701 SUSHI 7.
FT	DOMAIN 642 701 SUSHI 8.
FT	DISULFID 60 158 BY SIMILARITY.
FT	DISULFID 131 150 BY SIMILARITY.
FT	DISULFID 163 174 BY SIMILARITY.
FT	DISULFID 168 183 BY SIMILARITY.
FT	DISULFID 185 194 BY SIMILARITY.
FT	DISULFID 200 244 BY SIMILARITY.
FT	DISULFID 230 257 BY SIMILARITY.
FT	DISULFID 262 306 BY SIMILARITY.

CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch)	
CC	-----	
DR	EMBL: X63729; CAA45272.1; -	
DR	EMBL: X63728; CAA45271.1; -	
DR	PIR: S19852; S19852.	
DR	PIR: S20881; S20881.	
DR	HSP: P02833; 9ANT.	
DR	TRANSFAC: T02084; -	
DR	FLYBase: FBgn0003040; pb.	
DR	InterPro: IPR001827; Antennapedia.	
DR	InterPro: IPR001356; Homeobox.	
DR	Pfam: PF00046; homeobox; 1.	
DR	PRINTS: PR00025; ANTENNAPEDIA.	
DR	PRINTS: PR00024; HOMEBOX.	
DR	ProDom: PD000010; Homeobox; 1.	
DR	SMART: SM00389; Hox; 1.	
DR	SMART: PS00027; HOMEBOX_1; 1.	
DR	PROSITE: PS00032; ANTENNAPEDIA; 1.	
DR	PROSITE: PS00071; HOMEBOX_2; 1.	
DR	PROSITE: PS00071; Homeobox; Developmental protein; Nuclear protein;	
KW	Homeobox; DNA-binding; Developmental protein; Nuclear protein;	
KW	Alternative splicing.	
FT	DOMAIN 164	ANTP-TYPE HEXAPEPTIDE.
FT	DOMAIN 198	HOMEBOX.
FT	DNA_BIND 472	HOMEBOX.
FT	DOMAIN 535	GLN-RICH.
FT	DOMAIN 580	HIS/GLN-RICH.
FT	VARSPLIC 184	GDNSTT -> A (IN ISOFORM 2).
FT	VARSPLIC 189	TEFVEE -> K (IN ISOFORM 3).
FT	VARSPLIC 184	MISSING (IN ISOFORM 4).
FT	VARSPLIC 193	MISSING (IN ISOFORM 4).
SS	798 AA; 85559 MW; 886289FAF6C1758 CRC64;	
SS	Score 37; DB 1; Length 798;	

```

Matches 5; Conservative 5; Miscellaneous 5
Oy 1 QASIEGRYAPOP 13
|:::| |::|
Db 516 QQALDGEYLSPKR 528

RESULT 13
YP73.CAEEL STANDARD; PRT; 4385 AA.
ID YP73.CAEEL
AD Q09222;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 486.0 kDa protein B0228.3 in chromosome II.
DE B0228.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBL_TaxID=6239;
OX
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol NZ;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RA
RA -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
CC EMBL: U23168; AAC3807.1; -.
CC WormRep: B0228.3; CE01744.
CC Hypothetical protein.

```


DR EMBL: Y64359; CAA45709.1; -
 DR EMBL: D67027; BAA1043.1; -
 DR PIR: S14772; S14772.
 DR PIR: S19713; S19713.
 DR HSSP: P13449; 2AH7.
 DR InterPro: IPR003168; NHase_beta.
 DR Pfam: PF02211; NHase_beta; 1.
 KW Lyase.
 FT CONFLICT 7 T -> I (IN REF. 3).
 FT CONFLICT 26 MISSING (IN REF. 3).
 SQ SEQUENCE 229 AA; 26321 MW; 73270C13358D4671 CRC64;

Query Match 49.3%; Score 36; DB 1; length 229;
 Best Local Similarity 53.8%; Pred. No. 27;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QASIEGRYTAPOP 13
 | :||||| :|
 Db 99 QETLEGRYTDKRP 111

Search completed: January 13, 2003, 09:54:09
 Job time : 16.3333 secs


```

OM protein - protein search: 09:53:20 : Search engine updates/sec
                                (without alignments)
January 13, 2003, 33.750 Million cell
Run on:

```

sequence: RIOSUM62 Capext 0.5

Searched: 2000000 bits satisfying chosen parameters

```

Minimum DB seq length: 2000000
Minimum DB seq      Match 0%
Maximum DB seq      Match 100%

```

COMB. pep: *

```

Database :
1: /cqn2-6/ptodata/1/iaa/5A.COMB.pcp:*
2: /cqn2-6/ptodata/1/iaa/5A.COMB.pcp:*
3: /cqn2-6/ptodata/1/iaa/5A.COMB.pcp:*
Issued.patents.AA:*
/cqn2-6/ptodata/1/iaa/5B.COMB.pcp:*
/cqn2-6/ptodata/1/iaa/5A.COMB.pcp:*

```

- 4: /cgn2_-6/plodata/1/laa/backlinc
- 5: /cgn2_-6/plodata/1/laa/backlinc
- 6: /cgn2_-6/plodata/1/laa/backlinc

pred. no. than of the
score greater by analysis of
and is derived by analysis of

SUMMARIES

Result	Score	Query Match	length	DB	ID	Description
NO.						
1	38	66.7	1004	4	US-08-916-552-2	Sequence 2, Appli
2	36	63.2	532	4	US-09-322-478-18	Sequence 8, Appli
3	36	63.2	1802	4	US-08-533-205A-6	Sequence 6, Appli
4	35	61.4	834	4	US-08-583-562B-8	Sequence 8, Appli
5	35	61.4	2647	2	US-08-779-115-8	Sequence 8, Appli
6	35	61.4	2647	2	US-08-956-242-2	Sequence 2, Appli
7	34	59.6	6262	2	US-09-351-115-2	Sequence 4, Appli
8	34	59.6	626	3	US-08-910-925-4	Sequence 9, Appli
9	34	59.6	703	4	US-09-314-268-94	Sequence 8, Appli
10	33	57.8	35	4	US-09-109-100-9	Sequence 9, Appli
11	33	57.8	209	4	US-09-109-100-11	Sequence 11, Appli
12	33	57.8	209	4	US-09-109-100-12	Sequence 12, Appli
13	33	57.8	209	4	US-09-109-100-13	Sequence 13, Appli
14	33	57.8	209	4	US-09-109-100-14	Sequence 14, Appli
15	33	57.8	209	4	US-09-109-100-15	Sequence 15, Appli
16	33	57.8	209	4	US-09-109-100-16	Sequence 16, Appli
17	33	57.8	209	4	US-09-109-100-17	Sequence 17, Appli
18	33	57.8	209	4	US-09-109-100-18	Sequence 18, Appli
19	33	57.8	209	4	US-09-109-100-19	Sequence 19, Appli
20	33	57.8	209	4	US-09-109-100-20	Sequence 20, Appli
21	33	57.8	212	4	US-08-243-545-6	Sequence 6, Appli
22	33	57.8	235	1	US-08-993-662-6	Sequence 6, Appli
23	33	57.8	235	2	US-09-160-841-6	Sequence 6, Appli
24	33	57.8	235	4	US-09-109-100-1	Sequence 1, Appli
25	33	57.8	235	4	PCT-US94-05365-6	Sequence 6, Appli
26	33	57.8	251	2	US-08-592-214A-12	Sequence 12, Appli
27	33	57.8				

28	3	57.9	251	3	US-08-459-188-12
29	33	57.9	251	3	US-08-455-227-12
30	33	57.9	251	3	US-08-465-241-12
31	33	57.9	251	3	US-08-149-976-12
32	33	57.9	251	4	US-09-149-976-12
33	33	57.9	251	4	US-09-198-326-12
34	33	57.9	251	4	US-09-424-342-12
35	33	57.9	251	4	US-09-424-342-12
36	33	57.9	251	4	US-09-080-202-5
37	33	57.9	251	4	US-09-080-202-5
38	33	57.9	251	4	US-09-080-202-5
39	33	57.9	251	4	US-09-080-202-5
40	33	57.9	251	4	US-09-080-202-5
41	33	57.9	251	4	US-09-080-202-5
42	33	57.9	251	4	US-09-080-202-5
43	33	57.9	251	4	US-09-080-202-5
44	33	57.9	251	4	US-09-080-202-5
45	33	57.9	251	4	US-09-080-202-5

ALIGNMENTS

RESULT 1 US/08916352
US-08-916-352-2 Application
Sequence 2, 6166191

Patent No. 8-2-
GENERAL INFORMATION: CORPORATION
GENERAL INFORMATION: CHIRON POLYHOMEOT
APPLICANT: HUMAN SUPPRESSOR
INVENTION: TUMOR

TITLE OF INVENTION: 2
TITLE OF SEQUENCES: 2
NUMBER OF ADDRESSES:
CORPORATION

ADDRESS: 4560 HORTON STREET
CITY: CHIRON CORP
STATE: CHIRON CORP

SINCE: EMERYVILLE
CITY: CA
STATE: USA
COUNTRY:

COUNTRY: 94608
ZIP: 94608
COMPUTER READABLE TYPE:

MEDIUM: IBM PC Computer, Version #1.3
COMPUTER: IBM PC-DOS/MS-DOS
SYSTEM: PC-DOS/MS-DOS #1.0,
OPERATING: patentin Release

```

; SOFTWARE: FUSION
; CURRENT APPLICATION DATA: US/08/916, 352
; CURRENT APPLICATION NUMBER:
; APPLICATION
;

```

FILED DATE: 435
CLASSIFICATION: 435
AGENCY/AGENT INFORMATION: 435

ATTORNEY: POTTER, JANE
NAME: POTTER, JANE
REGISTRATION NUMBER: 33,332
REGISTRATION NUMBER: 1355
REGISTRATION: 1355
REGISTRATION: 1355

REFERENCE/COMMUNICATION INFORMATION
TELECOMMUNICATION 510-923-2707
TELEPHONE: 510-655-3542

TELEFAX: 510 833 2111
FOR SEQ ID NO: 2
INFORMATION FOR SEQ ID NO: 2
CHARACTERISTICS:
no acids

SEQUENCE: 1004 amino
LENGTH: amino acid
TYPE: single
PURITY: 100%

STRANDED: linear
TOPOLOGY: protein
TYPE: molecule

US-08-916-352-2

Query: Similarity
Best Local: Conservati
Matches: 7; 10

1 QASSYTAPQP 10
11:: 11111
11111TAPQP 450
QY

Db
441 QAL

US-09-554-941-2.rai

RESULT 2
US-09-322-478-8
Sequence 8, Application US/09322478
Patent No. 6331662
GENERAL INFORMATION:
APPLICANT: Wright, David A.
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
CURRENT APPLICATION NUMBER: US/09/322,478
EARLIER FILING DATE: 1999-05-28
EARLIER APPLICATION NUMBER: 60/087125
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 8
LENGTH: 532
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-8
Query Match
Best Local Similarity 63.2%; Score 36; DB 4; Length 532;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 QASSTYAPOP 10
DB 36 EGSSYSTSP 45
RESULT 3
US-09-322-478-18
Sequence 18, Application US/09322478
Patent No. 6331662
GENERAL INFORMATION:
APPLICANT: Wright, David A.
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
CURRENT APPLICATION NUMBER: US/09/322,478
EARLIER FILING DATE: 1999-05-28
EARLIER APPLICATION NUMBER: 60/087125
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 18
LENGTH: 1802
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-18
Query Match
Best Local Similarity 60.0%; Score 36; DB 4; Length 1802;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 QASSTYAPOP 10
DB 36 EGSSYSTSP 45
RESULT 4
US-08-539-205A-6
Sequence 6, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
CURRENT APPLICATION NUMBER: US/08/539,205A
EARLIER FILING DATE: 1999-05-28
EARLIER APPLICATION NUMBER: 60/087125
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 6
LENGTH: 532
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: plant
US-08-539-205A-6
Query Match
Best Local Similarity 61.4%; Score 35; DB 3; Length 834;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 QASSTYAPOP 11
DB 340 QPSYNSPKPO 350
RESULT 5
US-08-583-562B-8
Sequence 8, Application US/08583562B
Patent No. 5922570
GENERAL INFORMATION:
APPLICANT: Staunton, Donald
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF INVENTION: Binding
CORRESPONDENCE ADDRESSES: 36
ADDRESS: Marshall
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln R-DOS/MS-DOS
CURRENT APP. Patentln Release #1.0, Version #1.25
APPLICATION DATA:
FILING DATE: US/08/583,562B
CLASSIFICATION: 536
NAME: AGENT INFORMATION:
REGISTRATION NUMBER: Joseph A.
REFERENCE/DOCKET NUMBER: 38,659
TELECOMMUNICATION INFORMATION: 27866/33033

APPLICANT: Neisky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES: 6
ADDRESS: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln R-DOS/MS-DOS
CURRENT APP. Patentln Release #1.0, Version #1.30
APPLICATION DATA:
FILING DATE: US/08/539,205A
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELEPHONE: (617) 832-7000
TELEFAX: (617) 832-1000
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-539-205A-6
Query Match
Best Local Similarity 54.4%; Score 35; DB 3; Length 834;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 QASSTYAPOP 11
DB 340 QPSYNSPKPO 350
RESULT 5
US-08-583-562B-8
Sequence 8, Application US/08583562B
Patent No. 5922570
GENERAL INFORMATION:
APPLICANT: Staunton, Donald
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
NUMBER OF INVENTION: Binding
CORRESPONDENCE ADDRESSES: 36
ADDRESS: Marshall
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentln R-DOS/MS-DOS
CURRENT APP. Patentln Release #1.0, Version #1.25
APPLICATION DATA:
FILING DATE: US/08/583,562B
CLASSIFICATION: 536
NAME: AGENT INFORMATION:
REGISTRATION NUMBER: Joseph A.
REFERENCE/DOCKET NUMBER: 38,659
TELECOMMUNICATION INFORMATION: 27866/33033

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-583-562B-8

Query Match 61.4%; Score 35; DB 2; Length 2647;
Best Local Similarity 100.0%; Pred. NO. 1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YTAPOP 10
111111
DB 1712 YTAPOP 1717

RESULT 6
US-08-779-113-8
Sequence 8, Application US/08779113
Patent No. 5948891

GENERAL INFORMATION:
APPLICANT: Staunton, Donald E.
APPLICANT: Harris, Edith S.
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin
TITLE OF INVENTION: Binding
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,113
FILING DATE:

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 5948891and
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/33773
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 amino acids
TYPE: amino acid
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-779-113-8

Query Match 61.4%; Score 35; DB 2; Length 2647;
Best Local Similarity 100.0%; Pred. NO. 1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YTAPOP 10
111111
DB 1712 YTAPOP 1717

RESULT 7
US-08-956-242-2
Sequence 2, Application US/08956242C
Patent No. 5986081

GENERAL INFORMATION:
APPLICANT: Ganetzky, Barry S.
APPLICANT: Titus, Steven A.
TITLE OF INVENTION: Polynucleotides Encoding Herg-3
FILE REFERENCE: 960296.94550
CURRENT APPLICATION NUMBER: US/08/956,242C
CURRENT FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: unsure
LOCATION: (441)..(542)
OTHER INFORMATION: Unidentified at time of filing
US-08-956-242-2

Query Match 59.6%; Score 34; DB 2; Length 626;
Best Local Similarity 60.0%; Pred. NO. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASSTYAPPOQ 11
111111
DB 67 SSHYTFPPRQ 76

RESULT 8
US-09-351-215-2
Sequence 2, Application US/09351215
Patent No. 6087488

GENERAL INFORMATION:
APPLICANT: Ganetzky, Barry S.
APPLICANT: Titus, Steven A.
TITLE OF INVENTION: Polynucleotides Encoding Herg-3
FILE REFERENCE: 960296.94550
CURRENT APPLICATION NUMBER: US/09/351,215
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: 08/956,242
EARLIER FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: unsure
LOCATION: (441)..(542)
OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-2

Query Match 59.6%; Score 34; DB 3; Length 626;
Best Local Similarity 60.0%; Pred. NO. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASSTYAPPOQ 11
111111
DB 67 SSHYTFPPRQ 76

RESULT 9
US-08-910-925-4
Sequence 4, Application US/08910925
Patent No. 6162601

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/910,925

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0365 US

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 703 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1684843

US-08-910-925-4

Db 26 QTSHTAPQ 34

RESULT 11

Sequence 8, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 8

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-8

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 12

Sequence 9, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 9

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-9

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 60.0%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 13

Sequence 11, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 11

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-11

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 14

Sequence 12, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 12

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-12

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 15

Sequence 13, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 13

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-13

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 16

Sequence 14, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 14

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-14

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 17

Sequence 15, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 15

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-15

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 18

Sequence 16, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 16

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-16

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 19

Sequence 17, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 17

LENGTH: 209

TYPE: PRT

ORGANISM: Homo sapiens

US-09-109-100-17

Db 149 EATAPAPQ 158

Query Match

Best Local Similarity 57.9%; Score 33; DB 4; Length 209;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 QASSTAPQ 10

11111111

Db 149 EATAPAPQ 158

RESULT 20

Sequence 18, Application US/09109100C

Patent No. 6291661

GENERAL INFORMATION:

APPLICANT: Graddis, Thomas J.

APPLICANT: McGrew, Jeffrey T.

TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE

FILE REFERENCE: 03260.0028

CURRENT APPLICATION NUMBER: US/09/109,100C

CURRENT FILING DATE: 1998-07

LENGTH: 209
TYPE: PRT
ORGANISM: Homo sapiens
US-09-109-100-11

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

RESULT 14
US-09-109-100-12
; Sequence 12, Application US/09109100C
; Patent No. 6291661
; GENERAL INFORMATION:
; APPLICANT: Graddis, Thomas J.
; APPLICANT: McGrew, Jeffrey T.
; TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE
; FILE REFERENCE: 03260.0028
; CURRENT APPLICATION NUMBER: US/09/109,100C
; CURRENT FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 12
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-109-100-12

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

RESULT 15
US-09-109-100-13
; Sequence 13, Application US/09109100C
; Patent No. 6291661
; GENERAL INFORMATION:
; APPLICANT: Graddis, Thomas J.
; APPLICANT: McGrew, Jeffrey T.
; TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE
; FILE REFERENCE: 03260.0028
; CURRENT APPLICATION NUMBER: US/09/109,100C
; CURRENT FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 13
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-109-100-13

Query Match
Best Local Similarity 60.0%; Score 33; DB 4; Length 209;
Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
:|::|||||
Db 149 EATAPAPOP 158

Search completed: January 13, 2003, 09:56:24
Job time : 11.5897 secs

Mon Jan 13 10:35:13 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

January 13, 2003, 09:54:16 ; Search time 51.0513 Seconds
(without alignments)
4.180 Million cell updates/sec

Run on:

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQPO 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB_PEP:.*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB_PEP:.*
3: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB_PEP:.*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB_PEP:.*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB_PEP:.*
6: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB_PEP:.*
7: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB_PEP:.*
8: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB_PEP:.*
9: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB_PEP:.*
10: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB_PEP:.*
11: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB_PEP:.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB_PEP:.*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB_PEP:.*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB_PEP:.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	52	91.2	639	US-09-782-906-2	Sequence 2, Appl1
2	52	91.2	639	US-09-782-906-3	Sequence 3, Appl1
3	52	91.2	639	US-09-782-906-4	Sequence 4, Appl1
4	52	91.2	639	US-09-782-906-5	Sequence 5, Appl1
5	52	91.2	639	US-09-864-761-33417	Sequence 33417, A
6	52	91.2	639	US-09-864-761-33417	Sequence 324, App
7	52	91.2	639	US-09-864-761-33417	Sequence 160, App
8	52	91.2	639	US-09-864-761-33417	Sequence 8, Appl1
9	52	91.2	639	US-09-864-761-33417	Sequence 16, Appl
10	52	91.2	639	US-09-864-761-33417	Sequence 33417, A
11	52	91.2	639	US-09-864-761-33417	Sequence 324, App
12	52	91.2	639	US-09-864-761-33417	Sequence 160, App
13	52	91.2	639	US-09-864-761-33417	Sequence 8, Appl1
14	52	91.2	639	US-09-864-761-33417	Sequence 16, Appl
15	52	91.2	639	US-09-864-761-33417	Sequence 33417, A
16	52	91.2	639	US-09-864-761-33417	Sequence 324, App
17	52	91.2	639	US-09-864-761-33417	Sequence 160, App
18	52	91.2	639	US-09-864-761-33417	Sequence 8, Appl1
19	52	91.2	639	US-09-864-761-33417	Sequence 16, Appl

us-09-554-941-2.rapp

20	33	57.9	209	10	US-09-904-536-13	Sequence 13, Appl
21	33	57.9	209	10	US-09-904-536-14	Sequence 14, Appl
22	33	57.9	209	10	US-09-904-536-15	Sequence 15, Appl
23	33	57.9	209	10	US-09-904-536-16	Sequence 16, Appl
24	33	57.9	209	10	US-09-904-536-17	Sequence 17, Appl
25	33	57.9	209	10	US-09-904-536-18	Sequence 18, Appl
26	33	57.9	209	10	US-09-904-536-19	Sequence 19, Appl
27	33	57.9	209	10	US-09-904-536-20	Sequence 20, Appl
28	33	57.9	209	10	US-09-904-536-21	Sequence 21, Appl
29	33	57.9	209	10	US-09-904-536-22	Sequence 22, Appl
30	33	57.9	209	10	US-09-904-536-23	Sequence 23, Appl
31	33	57.9	209	10	US-09-904-536-24	Sequence 24, Appl
32	33	57.9	209	10	US-09-904-536-25	Sequence 25, Appl
33	33	57.9	209	10	US-09-904-536-26	Sequence 26, Appl
34	33	57.9	209	10	US-09-904-536-27	Sequence 27, Appl
35	33	57.9	209	10	US-09-904-536-28	Sequence 28, Appl
36	33	57.9	209	10	US-09-904-536-29	Sequence 29, Appl
37	33	57.9	209	10	US-09-904-536-30	Sequence 30, Appl
38	33	57.9	209	10	US-09-904-536-31	Sequence 31, Appl
39	33	57.9	209	10	US-09-904-536-32	Sequence 32, Appl
40	33	57.9	209	10	US-09-904-536-33	Sequence 33, Appl
41	33	57.9	209	10	US-09-904-536-34	Sequence 34, Appl
42	33	57.9	209	10	US-09-904-536-35	Sequence 35, Appl
43	33	57.9	209	10	US-09-904-536-36	Sequence 36, Appl
44	33	57.9	209	10	US-09-904-536-37	Sequence 37, Appl
45	33	57.9	209	10	US-09-904-536-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-782-906-2
Sequence 2, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.
APPLICANT: Murphy, Dennis J.
APPLICANT: Welfa III, Anthony M.
APPLICANT: Bylani, Edward J.
APPLICANT: Coleman, William Galactose Oxidase, Nucleic Acid Encoding Same, And Met
TITLE OF INVENTION: Using Same
FILE REFERENCE: HER-0040
CURRENT FILING DATE: 2001-02-14
PRIORITY FILING DATE: 2000-02-25
PRIORITY FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 639
TYPE: PRT
ORGANISM: Dactylium dendroideus
US-09-782-906-2
Query Match
Best Local Similarity
Matches
DB 147 QASSTAPQPO 156
91.2% Score 52; DB 10; Length 639;
100.0% Pred No. 0.14; 0; Indels 0; Gaps 0;
Mismatches 0;
RESULT 2
US-09-782-906-3
Sequence 3, Application US/09782906
Patent No. US20010051369A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Rittenhouse Pruss, Jennifer L.
APPLICANT: Murphy, Dennis J.
APPLICANT: Welfa III, Anthony M.
APPLICANT: Bylani, Edward J.
APPLICANT: Coleman, William Galactose Oxidase, Nucleic Acid Encoding Same, And Met
TITLE OF INVENTION: Using Same
FILE REFERENCE: HER-0040
CURRENT FILING DATE: 2001-02-14
PRIORITY FILING DATE: 2000-02-25
PRIORITY FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 639
TYPE: PRT
ORGANISM: Dactylium dendroideus
US-09-782-906-3
Query Match
Best Local Similarity
Matches
DB 147 QASSTAPQPO 156
91.2% Score 52; DB 10; Length 639;
100.0% Pred No. 0.14; 0; Indels 0; Gaps 0;
Mismatches 0;

```

; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Method
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-3

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 3
US-09-782-906-4
; Sequence 4, Application US/09782906
; Patent No. US20010051369A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Rittenhouse Pruss, Jennifer L.
; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Method
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-4

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 4
US-09-782-906-5
; Sequence 5, Application US/09782906
; Patent No. US20010051369A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon

```

```

; APPLICANT: Rittenhouse Pruss, Jennifer L.
; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Method
; FILE REFERENCE: HRP-0040
; CURRENT APPLICATION NUMBER: US/09/782,906
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-5

```

```

Query Match
Best Local Similarity 91.28; Score 52; DB 10; Length 639;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156

```

```

RESULT 5
US-09-864-761-33417
; Sequence 33417, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DEIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rapb

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-28
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33417
LENGTH: 223
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: MAP TO AC006581.16
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 41
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 23
OTHER INFORMATION: EXPRESSED IN RETA, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 71
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 15
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 29
OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 29
OTHER INFORMATION: EST_HUMAN HIT: AW959289.1, EVALUO 9.00e-15
US-09-864-761-33417

Query Match Similarity 66.7%; Score 38; DB 10; Length 223;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 6
US-09-893-737-324
Sequence 324, Application US/09893737
Patent No. US20020110855A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Plesnell, Scott R.
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
FILE REFERENCE: 00-41
CURRENT APPLICATION NUMBER: US/09/893,737
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,446
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 329
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 324
LENGTH: 295
TYPE: PRT
ORGANISM: Homo sapiens
US-09-893-737-324

Query Match Similarity 64.9%; Score 37; DB 10; Length 295;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
US-09-729-674-160
Sequence 160, Application US/09729674
Patent No. US20010039335A1
GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Reckel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 160
LENGTH: 423
TYPE: PRT
ORGANISM: Homo sapiens
US-09-729-674-160

Query Match Similarity 64.9%; Score 37; DB 10; Length 423;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
US-09-965-553-8
Sequence 8, Application US/09965553
Patent No. US20020112259A1
GENERAL INFORMATION:
APPLICANT: Wright, David A.
APPLICANT: Voytas, Daniel F.
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
FILE REFERENCE: P-1065 ISUR Plant Retroelement
CURRENT APPLICATION NUMBER: US/09/965,553
CURRENT FILING DATE: 2001/09/322,478
PRIOR APPLICATION NUMBER: 1999-05-28/087125
PRIOR FILING DATE: 1998-05-25
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 532
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: plant
OTHER INFORMATION: retroelement sequence
US-09-965-553-8

Query Match Similarity 63.2%; Score 36; DB 10; Length 532;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
US-09-965-553-8
Sequence 8, Application US/09965553
Patent No. US20020112259A1
GENERAL INFORMATION:
APPLICANT: Wright, David A.
APPLICANT: Voytas, Daniel F.
TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
FILE REFERENCE: P-1065 ISUR Plant Retroelement
CURRENT APPLICATION NUMBER: US/09/965,553
CURRENT FILING DATE: 2001/09/322,478
PRIOR APPLICATION NUMBER: 1999-05-28/087125
PRIOR FILING DATE: 1998-05-25
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 532
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: plant
OTHER INFORMATION: retroelement sequence
US-09-965-553-8

Query Match Similarity 63.2%; Score 36; DB 10; Length 532;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rapb

Page 4

```
us-09-965-553-18
; Sequence 18, Application US/09965553
; Patent No. US20020112259A1
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; FILE REFERENCE: P-1065 ISUP Plant Retroelement
; CURRENT FILING DATE: 2001-09-27
; PRIOR FILING DATE: 1999-05-28/322,478
; PRIOR APPLICATION NUMBER: 09/322,478
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO: 18
; LENGTH: 1802
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Plant
; US-09-965-553-18

Query Match
Best Local Similarity 63.2%; Score 36; DB 10; Length 1802;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 QASSTAPOP 10
DB 36 EGSSYSPSP 45

RESULT 10
; Sequence 35691, Application US/09864761
; Patent No. US2002004675A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Neomula-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: CB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
us-09-738-626-6745
; Sequence 6745, Application US/09738626
; Patent No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent Ver. 3.0
; SEQ ID NO: 6745
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6745

Query Match
Best Local Similarity 77.8%; Score 35; DB 10; Length 53;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ASSYTAPOP 10
DB 44 ASSYTAPOP 52

RESULT 11
; Sequence 6745, Application US/09738626
; Patent No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent Ver. 3.0
; SEQ ID NO: 6745
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-6745
```

Query Match 61.4%; Score 35; DB 9; Length 301;
Best Local Similarity 87.5%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ASSYTAPO 9
11111111
Db 38 ASSYTAPO 45

RESULT 12
US-09-764-864-1515
; Sequence 1515, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1515
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1515

Query Match 59.6%; Score 34; DB 10; Length 294;
Best Local Similarity 54.5%; Pred. No. 71;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 QASSTAPPO 11
11111111
Db 181 QSSSPTSPEPR 191

RESULT 13
US-09-764-864-947
; Sequence 947, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 947
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-947

Query Match 59.6%; Score 34; DB 10; Length 323;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 QASSTAPPO 11
11111111
Db 210 QSSSPTSPEPR 220

RESULT 14
US-09-764-864-1095
; Sequence 1095, Application US/09764864
; Patent No. US20020132753A1

GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1095
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1095

Query Match 59.6%; Score 34; DB 10; Length 323;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 QASSTAPPO 11
11111111
Db 210 QSSSPTSPEPR 220

RESULT 15
US-09-738-626-5433
; Sequence 5433, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentln ver. 3.0
; SEQ ID NO 5433
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5433

Query Match 59.6%; Score 34; DB 9; Length 407;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 YRAPPO 11
11111111
Db 254 YQAPPO 260

Search completed: January 13, 2003, 10:02:49
Job time : 51.0513 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model
Run on: January 13, 2003, 09:53:15
(without alignments)
58.391 Million cell updates/sec

```

Title: US-09-554-941-2
perfect score: 57
sequence: 1 QASSYTAPOQ 11

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 locs
 908470

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

```

post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

A: Geneseq - 101002: *
1: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1960. DAT: *
2: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1961. DAT: *
3: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1962. DAT: *
4: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1963. DAT: *
5: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1964. DAT: *
6: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1965. DAT: *
7: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1966. DAT: *
8: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1967. DAT: *
9: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1968. DAT: *
10: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1969. DAT: *
11: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1970. DAT: *
12: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1971. DAT: *
13: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1972. DAT: *
14: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1973. DAT: *
15: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1995. DAT: *
16: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1996. DAT: *
17: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1997. DAT: *
18: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1998. DAT: *
19: /SID52/gcgdata/geneseq/geneseqp -emb1/AA1999. DAT: *
20: /SID52/gcgdata/geneseq/geneseqp -emb1/AA2000. DAT: *
21: /SID52/gcgdata/geneseq/geneseqp -emb1/AA2001. DAT: *
22: /SID52/gcgdata/geneseq/geneseqp -emb1/AA2002. DAT: *
23:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	100.0	11	20	AAV22919	Galactose oxidase
2	52	91.2	639	22	AAE12925	Dactylium dendroid
3	52	91.2	639	22	AAE12926	Dactylium dendroid
4	52	91.2	639	22	AAE12927	Dactylium dendroid
5	52	91.2	639	22	AAE12928	Dactylium dendroid
6	52	91.2	639	22	AAE12929	Dactylium dendroid
7	52	91.2	639	22	AAE12930	Dactylium dendroid
8	52	91.2	639	22	AAE12931	Dactylium dendroid
9	52	91.2	639	22	AAE12932	Dactylium dendroid
10	52	91.2	639	22	AAE12933	Dactylium dendroid

11	52	91.2	639	22	AAE12934	Dactylium dendroid
12	52	91.2	639	22	AAE12935	Dactylium dendroid
13	52	91.2	639	22	AAE12937	Dactylium dendroid
14	52	91.2	639	22	AAE12939	Dactylium dendroid
15	52	91.2	639	22	AAE12940	Dactylium dendroid
16	52	91.2	639	22	AAE12942	Dactylium dendroid
17	52	91.2	639	22	AAE12944	Dactylium dendroid
18	52	91.2	639	22	AAE12945	Dactylium dendroid
19	52	91.2	639	22	AAE12946	Dactylium dendroid
20	52	91.2	639	22	AAE12958	Dactylium dendroid
21	52	91.2	639	22	AAE12959	Dactylium dendroid
22	52	91.2	639	22	AAE12960	Dactylium dendroid
23	52	91.2	639	22	AAE12961	Dactylium dendroid
24	52	91.2	639	22	AAE12962	Dactylium dendroid
25	52	91.2	639	22	AAE12963	Dactylium dendroid
26	52	91.2	639	22	AAE12964	Dactylium dendroid
27	52	91.2	639	22	AAE12965	Dactylium dendroid
28	52	91.2	639	22	AAE12966	Dactylium dendroid
29	52	91.2	639	22	AAE12967	Dactylium dendroid
30	52	91.2	639	22	AAE12968	Dactylium dendroid
31	52	91.2	639	22	AAE12968	Dactylium dendroid
32	52	91.2	639	23	ABB05136	Galactose oxidase
33	52	91.2	639	23	ABB05137	Galactose oxidase
34	52	91.2	639	23	ABB05138	Galactose oxidase
35	52	91.2	639	23	ABB05139	Galactose oxidase
36	52	91.2	639	23	ABB05140	Galactose oxidase
37	52	91.2	639	23	ABB05141	Galactose oxidase
38	52	91.2	639	23	ABB05142	Galactose oxidase
39	52	91.2	639	23	ABB05143	Galactose oxidase
40	52	91.2	639	23	ABB05144	Galactose oxidase
41	52	91.2	639	23	ABB05145	Galactose oxidase
42	52	91.2	639	23	ABB05146	Galactose oxidase
43	52	91.2	639	23	ABB05147	Galactose oxidase
44	52	91.2	680	22	AA653582	Amino acid sequenc
45	52	91.2	680	22	AA653582	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAY22919 standard: peptide; 11 AA

AC AAY22919

XX	19-AUG-1999 (first entry)
DT	

DE Galactose oxidase Liner 2000

XX peptide linker; fusion protein; pathogen resistance; thaliana; kw inhibitor; Arabidopsis

parasitic nematode

05 Synthetic.

PN WO9928484-A1

AA
PD
10-JUN-1999

XX
PF 01-DEC-1998; 98WO-EP07792.

XX 03-DEC-1997; 97GB-0025556
PR

XX
72 (NOVS) NOVARTIS AG.

XX
McPherson MJ,
Urwin PE

XX
1000-385387/32

inhibitor fusion proteins

XX
P1
XX
33: 3000: English

PS

XX The present sequence represents a peptide linker used in the fusion
 CC proteins of the invention. The specification describes a method for
 CC improving pathogen resistance or tolerance of a plant. The method
 CC comprises transformation of the plant with a transgene encoding a
 CC fusion protein which comprises two or more protein or domains that
 CC are capable of improving pathogen resistance or tolerance when
 CC expressed on their own. Specifically, two distinct protease
 CC inhibitors are co-delivered, as a fusion, to Arabidopsis thaliana. The
 CC proteins or domains are connected by a peptide linker. The method is
 CC used to improve pathogen resistance or tolerance of a plant and its
 CC descendants, especially against parasitic nematode attack.

SO Sequence 11 AA;

Query Match 100.0%; Score 57; DB 20; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QASSTYAPQPQ 11
 Db 1 QASSTYAPQPQ 11

RESULT 2
 AAE12925
 ID AAE12925 standard; Protein; 639 AA.
 AC AAE12925;
 XX
 DT 15-JAN-2002 (first entry)

DE Dactylium dendroides wild type galactose oxidase (GO) mature protein.
 XX
 KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque.
 OS Dactylium dendroides.
 XX
 PN W0200162938-A2.
 PD 30-AUG-2001.

PF 21-FEB-2001; 2001WO-US05732.
 PR 25-FEB-2000; 2000US-185001P.
 PR 14-FEB-2001; 2001US-0782906.
 XX
 PA (HERC) HERCULES INC.

PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 PI Coleman WJ;
 DR WPI: 2001-648219/74.
 DR N-PSDB; AAD21273.

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 PT acid -
 XX
 PS
 XX

Example 1; Page 10; 65pp; English.

CC The invention relates to variant galactose oxidase (vGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schliff's reagent for early detection, pulp
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.

CC The present sequence is Dactylium dendroides wild type galactose
 CC oxidase (GO) mature protein.
 CC
 XX
 SO Sequence 639 AA;

Query Match 91.2%; Score 52; DB 22; Length 639;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QASSTYAPQP 10
 Db 147 QASSTYAPQP 156

RESULT 3
 AAE12926
 ID AAE12926 standard; Protein; 639 AA.
 AC AAE12926;
 XX
 DT 15-JAN-2002 (first entry)

DE Dactylium dendroides galactose oxidase GO8-1 variant.
 XX
 KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
 KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
 XX muten; variant.
 OS Synthetic.
 OS Dactylium dendroides.
 XX
 PN W0200162938-A2.

PD 30-AUG-2001.
 PF 21-FEB-2001; 2001WO-US05732.
 PR 25-FEB-2000; 2000US-185001P.
 PR 14-FEB-2001; 2001US-0782906.
 XX
 PA (HERC) HERCULES INC.

PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
 PI Coleman WJ;
 DR WPI: 2001-648219/74.
 DR N-PSDB; AAD21274.

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 PT acid -
 XX
 PS
 XX

Claim 10; Page 45; 65pp; English.

CC The invention relates to variant galactose oxidase (vGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schliff's reagent for early detection, pulp
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.

Mon Jan 13 10:35:13 2003

us-09-554-941-2.rag

Page 3

The present sequence is Dactylium dendroides galactose oxidase (GO)

CC variant.

Sequence 639 AA: 91.2%; Score 52; DB 22; Length 639; Gaps 0;

Query Match 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;
Best Local Similarity 100.0%; 0; Mismatches 0; Gaps 0;
Matches 10; Conservative

OY 1 QASSTAPOP 10
Db 147 QASSTAPOP 156

RESULT 4

AAE12927 standard; Protein: 639 AA.

AC AAE12927;

DT 15-JAN-2002 (first entry)

Dactylium dendroides galactose oxidase GOB-1H3A variant;
Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
pulp biobleaching; colorectal adenocarcinoma; dental plaque; variant;

KW mutant; muten.
KW mutant; muten.
OS Dactylium dendroides.
OS Synthetic.

Location/Qualifiers

Key 63 /note= "Wild type Gln substituted with Lys"

Misc-difference 195 /note= "Wild type Gly substituted with Ala"

Misc-difference 383 /note= "Wild type Cys substituted with Ser"

Misc-difference 436 /note= "Wild type Tyr substituted with His"

Misc-difference 494 /note= "Wild type Val substituted with Ala"

MO200162938-A2.
30-AUG-2001.
21-FEB-2001; 2001MO-US05732.
25-FEB-2000; 2000US-185001P.
14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.
Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
Coleman WJ;
WPI: 2001-648219/74.
N-PSDB: AAD21275.
Novel variant galactose oxidase enzyme for oxidizing a compound such as
quar which is useful in paper making, has at least one substituted amino
aldehydes and pulp biobleaching, has at least one substituted amino
acid

Claim 10: Page 46; 65pp; English.

The invention relates to variant galactose oxidase (VGO) proteins and
their cDNA molecules. Galactose oxidases are useful for oxidizing
D-galactose side chains of many complex compounds such as guar, major
glycolipid of human red cells. GO is useful in paper making processes,
generation of H2O2 in situ, enzymatic synthesis of aldehydes and
biobleaching, the use of GO-Schiff's reagent for early detection and

prognosis in human colorectal adenocarcinoma and use of GO-glucan
binding domain fusion proteins as targeting inhibitors of dental
bacteria. GO is useful as a research tool for identifying
characterisation and purification of interacting regulatory proteins.
The present sequence is Dactylium dendroides galactose oxidase (GO)

Sequence 639 AA: 91.2%; Score 52; DB 22; Length 639; Gaps 0;

Query Match 100.0%; Pred. No. 0.82; 0; Indels 0; Gaps 0;
Best Local Similarity 100.0%; 0; Mismatches 0; Gaps 0;
Matches 10; Conservative

OY 1 QASSTAPOP 10
Db 147 QASSTAPOP 156

RESULT 5

AAE12928 standard; Protein: 639 AA.

AC AAE12928;

DT 15-JAN-2002 (first entry)

Dactylium dendroides galactose oxidase (GO) 7.3.2 variant;
Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
pulp biobleaching; colorectal adenocarcinoma; dental plaque; variant;

KW mutant; muten.
KW mutant; muten.
OS Dactylium dendroides.
OS Synthetic.

Location/Qualifiers

Key 248 /note= "Wild type Lys substituted with Glu"

Misc-difference 352 /note= "Wild type Thr substituted with Ser"

Misc-difference 366 /note= "Wild type Lys substituted with Arg"

Misc-difference 383 /note= "Wild type Cys substituted with Ser"

Misc-difference 436 /note= "Wild type Tyr substituted with His"

MO200162938-A2.
30-AUG-2001.
21-FEB-2001; 2001MO-US05732.
25-FEB-2000; 2000US-185001P.
14-FEB-2001; 2001US-0782906.

(HERC) HERCULES INC.
Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
Coleman WJ;
WPI: 2001-648219/74.
N-PSDB: AAD21276.
Novel variant galactose oxidase enzyme for oxidizing a compound such as
quar which is useful in paper making, has at least one substituted amino
aldehydes and pulp biobleaching, has at least one substituted amino
acid

Claim 10: Page 46; 65pp; English.

The invention relates to variant galactose oxidase (VGO) proteins and
their cDNA molecules. Galactose oxidases are useful for oxidizing
D-galactose side chains of many complex compounds such as guar, major
glycolipid of human red cells. GO is useful in paper making processes,
generation of H2O2 in situ, enzymatic synthesis of aldehydes and
biobleaching, the use of GO-Schiff's reagent for early detection and

15-JAN-2002 (first entry)

Dactylium dendroides galac

muteln; variant.

Location	Year	Key
Mississippi	1990	1
Mississippi	1991	2
Mississippi	1992	3
Mississippi	1993	4
Mississippi	1994	5
Mississippi	1995	6
Mississippi	1996	7
Mississippi	1997	8
Mississippi	1998	9
Mississippi	1999	10
Mississippi	2000	11
Mississippi	2001	12
Mississippi	2002	13
Mississippi	2003	14
Mississippi	2004	15
Mississippi	2005	16
Mississippi	2006	17
Mississippi	2007	18
Mississippi	2008	19
Mississippi	2009	20
Mississippi	2010	21
Mississippi	2011	22
Mississippi	2012	23
Mississippi	2013	24
Mississippi	2014	25
Mississippi	2015	26
Mississippi	2016	27
Mississippi	2017	28
Mississippi	2018	29
Mississippi	2019	30
Mississippi	2020	31
Mississippi	2021	32
Mississippi	2022	33
Mississippi	2023	34
Mississippi	2024	35
Mississippi	2025	36
Mississippi	2026	37
Mississippi	2027	38
Mississippi	2028	39
Mississippi	2029	40
Mississippi	2030	41
Mississippi	2031	42
Mississippi	2032	43
Mississippi	2033	44
Mississippi	2034	45
Mississippi	2035	46
Mississippi	2036	47
Mississippi	2037	48
Mississippi	2038	49
Mississippi	2039	50
Mississippi	2040	51
Mississippi	2041	52
Mississippi	2042	53
Mississippi	2043	54
Mississippi	2044	55
Mississippi	2045	56
Mississippi	2046	57
Mississippi	2047	58
Mississippi	2048	59
Mississippi	2049	60
Mississippi	2050	61
Mississippi	2051	62
Mississippi	2052	63
Mississippi	2053	64
Mississippi	2054	65
Mississippi	2055	66
Mississippi	2056	67
Mississippi	2057	68
Mississippi	2058	69
Mississippi	2059	70
Mississippi	2060	71
Mississippi	2061	72
Mississippi	2062	73
Mississippi	2063	74
Mississippi	2064	75
Mississippi	2065	76
Mississippi	2066	77
Mississippi	2067	78
Mississippi	2068	79
Mississippi	2069	80
Mississippi	2070	81
Mississippi	2071	82
Mississippi	2072	83
Mississippi	2073	84
Mississippi	2074	85
Mississippi	2075	86
Mississippi	2076	87
Mississippi	2077	88
Mississippi	2078	89
Mississippi	2079	90
Mississippi	2080	91
Mississippi	2081	92
Mississippi	2082	93
Mississippi	2083	94
Mississippi	2084	95
Mississippi	2085	96
Mississippi	2086	97
Mississippi	2087	98
Mississippi	2088	99
Mississippi	2089	100
Mississippi	2090	101
Mississippi	2091	102
Mississippi	2092	103
Mississippi	2093	104
Mississippi	2094	105
Mississippi	2095	106
Mississippi	2096	107
Mississippi	2097	108
Mississippi	2098	109
Mississippi	2099	110
Mississippi	2100	111
Mississippi	2101	112
Mississippi	2102	113
Mississippi	2103	114
Mississippi	2104	115
Mississippi	2105	116
Mississippi	2106	117
Mississippi	2107	118
Mississippi	2108	119
Mississippi	2109	120
Mississippi	2110	121
Mississippi	2111	122
Mississippi		

misc-difference	436	Wild type C
/notes		

30-AUG-2001.

25-FEB-2000; 2000US-185001P.
14-FEB-2001; 2001US-185001P.

Maffia AM,
Coleman WT,
Delagrave

Novel variant galactose oxidase

The invention relates to:

...their CDNA molecules to variant galactose oxidase (vGO) proteins and -galactose side chains of many complex compounds such as guar, major

Query Match	91.2%;	Score 52;	DB 22;	Length 639;
Best Local Similarity	100.0%;	Pred. No.	0.82;	
Matches 10;	Conservative	0.00;		

RESULT 2

15-7200 800-
X
X

Dactylium dendroides galac

colorectal adenocarcinoma; dental plaque; mutation; variant.

Key

reference 383 type gin substituted with Lys

494 type Tyr substituted with Asn"

20 AUG-2001.

14-FEB-2001; 2001US-0782906.

Maffia AM,
Coleman WJ;
Delagrave

novel variant galactose oxidase enzyme, which is

amino acid, is useful in paper making, for enzymatic synthesis of aldehydes and pulp bleaching, has at least one substituted amino

XX The invention relates to variant galactose oxidase (GO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156
|||||
AAEI2931 standard; Protein; 639 AA.
XX
AC AAEI2931:
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase variant, C383S/Y436X/Y494A/Q63K.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW mutcin; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 63 /note= "Wild type Gln substituted with Lys"
FT MISC-difference 383 /note= "Wild type Cys substituted with Ser"
FT MISC-difference 436 /label= Asn, His
FT MISC-difference 494 /note= "Wild type Val substituted with Ala"
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
XX WPI; 2001-648219/74.
XX
DR Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of

PT aldehydes and pulp biobleaching, has at least one substituted amino
PI acid -
XX
PS Claim 18; Page -: 65pp; English.
XX
XX The invention relates to variant galactose oxidase (GO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTYAPOP 10
Db 147 QASSTYAPOP 156
|||||
AAEI2932 standard; Protein; 639 AA.
XX
AC AAEI2932:
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase variant, C383S/Y436H.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW mutcin; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 383 /note= "Wild type Cys substituted with Ser"
FT MISC-difference 436 /note= "Wild type Tyr substituted with His"
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
XX WPI; 2001-648219/74.
XX
DR Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of

PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -
XX
PS Claim 19; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (vGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTAPQP 10
IIIIIIIIII
DB 147 QASSTAPQP 156
RESULT 10
AAE12933
ID AAE12933 standard; Protein; 639 AA.
XX
AC AAE12933;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.05h1b variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM mutein; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 115 /note= "Wild type Asn substituted with His"
FT
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

XX
PS Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (vGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTAPQP 10
IIIIIIIIII
DB 147 QASSTAPQP 156
RESULT 11
AAE12934
ID AAE12934 standard; Protein; 639 AA.
XX
AC AAE12934;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.05h1c variant.
XX
KM Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KM pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KM mutein; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 195 /note= "Wild type Gly substituted with Glu"
FT
FT Misc-difference 553 /note= "Wild type Ser substituted with Cys"
FT
XX
PN WO200162938-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

```
XX
PS Example 1; Page -: 65pp; English.
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their CDNA molecules. Galactose oxidases are useful for oxidising
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H202 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
XX
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYRPOP 10
Db 147 QASSTYRPOP 156
|||||
|
RESULT 12
AAE12935 standard; Protein; 639 AA.
XX
AC AAE12935;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO.1h1c variant.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Gly substituted with Arg"
FT WO200162938-A2.
XX
PN 30-AUG-2001.
XX
PD 21-FEB-2001; 2001MO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC ) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid
XX
PS Example 1; Page -: 65pp; English.
```

```
XX
CC The invention relates to variant galactose oxidase (VGO) proteins and
CC their CDNA molecules. Galactose oxidases are useful for oxidising
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H202 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.
XX
SQ Sequence 639 AA;
XX
Query Match 91.2%; Score 52; DB 22; Length 639;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSTYRPOP 10
Db 147 QASSTYRPOP 156
|||||
|
RESULT 13
AAE12936 standard; Protein; 639 AA.
XX
AC AAE12936;
XX
DT 15-JAN-2002 (first entry)
XX
DE Dactylium dendroides galactose oxidase GO8.1h1a variant.
XX
KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;
KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;
KW muten; variant.
XX
OS Dactylium dendroides.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 238 /note= "Wild type Gln substituted with Leu"
FT MISC-difference 342 /note= "Wild type Lys substituted with Glu"
FT MISC-difference 383 /note= "Wild type Cys substituted with Ser"
FT MISC-difference 436 /note= "Wild type Tyr substituted with His"
FT MISC-difference 494 /note= "Wild type Val substituted with Ala"
FT WO200162938-A2.
XX
PN 30-AUG-2001.
XX
PD 21-FEB-2001; 2001MO-US05732.
XX
PR 25-FEB-2000; 2000US-185001P.
PR 14-FEB-2001; 2001US-0782906.
XX
PA (HERC ) HERCULES INC.
XX
PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;
PI Coleman WJ;
XX
DR WPI; 2001-648219/74.
XX
```

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

PS Example 1; Page -; 65pp; English.

XX The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.

XX Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;

Best Local Similarity 100.0%; Pred. No. 0.82;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QASSTAPOP 10
|||||

DB 147 QASSTAPOP 156

RESULT 14

AAEL2937

ID AAEL2937 standard; Protein: 639 AA.

XX AAEL2937;

DT 15-JAN-2002 (first entry)

DE Dactylium dendroides galactose oxidase G08-1h2A variant.

KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;

KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;

KW muten; variant.

OS Dactylium dendroides.

OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 427 /note= "Wild type Asn substituted with Thr"

FT MISC-difference 436 /note= "Wild type Tyr substituted with His"

FT MISC-difference 494 /note= "Wild type Val substituted with Ala"

PN W0200162938-A2.

PD 30-AUG-2001.

PF 21-FEB-2001; 2001WO-US05732.

PR 25-FEB-2000; 2000US-185001P.

PR 14-FEB-2001; 2001US-0782906.

PA (HERC) HERCULES INC.

PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;

PI Coleman WJ;

XX

DR WPI; 2001-648219/74.

XX Novel variant galactose oxidase enzyme for oxidizing a compound such as
PT guar which is useful in paper making, for enzymatic synthesis of
PT aldehydes and pulp biobleaching, has at least one substituted amino
PT acid -

PS Example 1; Page -; 65pp; English.

XX The invention relates to variant galactose oxidase (VGO) proteins and
CC their cDNA molecules. Galactose oxidases are useful for oxidizing
CC D-galactose side chains of many complex compounds such as guar, major
CC glycolipid of human red cells. GO is useful in paper making processes,
CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
CC biobleaching, the use of GO-Schiff's reagent for early detection and
CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
CC binding domain fusion proteins as targeting inhibitors of dental
CC plaque bacteria. GO is useful as a research tool for identification,
CC characterisation and purification of interacting regulatory proteins.
CC The present sequence is Dactylium dendroides galactose oxidase (GO)
CC variant.
CC Note: The present sequence is not shown in the specification but is
CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
CC shown in page 10 of the specification.

XX Sequence 639 AA:

Query Match 91.2%; Score 52; DB 22; Length 639;

Best Local Similarity 100.0%; Pred. No. 0.82;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QASSTAPOP 10
|||||

DB 147 QASSTAPOP 156

RESULT 15

AAEL2939

ID AAEL2939 standard; Protein: 639 AA.

XX AAEL2939;

DT 15-JAN-2002 (first entry)

DE Dactylium dendroides galactose oxidase G08-1h4A variant.

KW Galactose oxidase; GO; paper making; aldehyde synthesis; detection;

KW pulp biobleaching; colorectal adenocarcinoma; dental plaque; mutant;

KW muten; variant.

OS Dactylium dendroides.

OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 63 /note= "Wild type Gln substituted with Lys"

PN W0200162938-A2.

PD 30-AUG-2001.

PF 21-FEB-2001; 2001WO-US05732.

PR 25-FEB-2000; 2000US-185001P.

PR 14-FEB-2001; 2001US-0782906.

PA (HERC) HERCULES INC.

PI Maffia AM, Delagrave S, Murphy DJ, Rittenhouse Pruss J, Bylina E;

PI Coleman WJ;

XX WPI; 2001-648219/74.

PT Novel variant galactose oxidase enzyme for oxidizing a compound such as
 PT guar which is useful in paper making, for enzymatic synthesis of
 PT aldehydes and pulp biobleaching, has at least one substituted amino
 PT acid

XX
 PS Example 1; Page -; 65pp; English.

XX
 CC The invention relates to variant galactose oxidase (VGO) proteins and
 CC their cDNA molecules. Galactose oxidases are useful for oxidizing
 CC D-galactose side chains of many complex compounds such as guar, major
 CC glycolipid of human red cells. GO is useful in paper making processes,
 CC generation of H2O2 in situ, enzymatic synthesis of aldehydes, pulp
 CC biobleaching, the use of GO-Schiff's reagent for early detection and
 CC prognosis in human colorectal adenocarcinoma and use of GO-glucan
 CC binding domain fusion proteins as targeting inhibitors of dental
 CC plaque bacteria. GO is useful as a research tool for identification,
 CC characterisation and purification of interacting regulatory proteins.
 CC The present sequence is Dactylium dendroides galactose oxidase (GO)
 CC variant.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from Dactylium dendroides galactose oxidase SEQ ID NO: 2
 CC shown in page 10 of the specification.
 XX

SQ Sequence 639 AA;

Query Match 91.2%; Score 52; DB 22; Length 639;
 Best Local Similarity 100.0%; Pred. NO. 0.82;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTAPOP 10
 |||||
 DB 147 QASSTAPOP 156

Search completed: January 13, 2003, 09:55:41
 Job time : 26.1026 secs

Mon Jan 13 10:35:14 2003

us-09-554-941-2.rspt

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 : Search time 41.7436 Seconds
(without alignments)
54.296 Million cell updates/sec

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQPO 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	70.2	533	5 Q8WP81	Q8WP81 caenorhabdi
2	39	68.4	224	16 Q9RU88	Q9RU88 delinococcus
3	39	68.4	587	13 Q9BSL7	Q9BSL7 xenopus lae
4	39	68.4	592	13 Q9BSL6	Q9BSL6 xenopus lae
5	39	68.4	906	3 Q9P6N4	Q9P6N4 schizosacch
6	38	66.7	124	12 Q9OH57	Q9OH57 gallid herp
7	38	66.7	295	10 Q949H3	Q949H3 nevaya brasi
8	38	66.7	859	11 Q9YDM6	Q9YDM6 mus musculu
9	38	66.7	1004	4 P78364	P78364 homo sapien
10	37	64.9	90	16 Q9RWZ8	Q9RWZ8 delinococcus
11	37	64.9	112	16 Q8YBW7	Q8YBW7 brucea me
12	37	64.9	422	3 Q9YU44	Q9YU44 sorghum bic
13	37	64.9	442	10 Q9XE77	Q9XE77 caenorhabdi
14	37	64.9	583	5 Q61711	Q61711 caenorhabdi
15	37	64.9	1465	3 Q9P5N0	Q9P5N0 schizosacch
16	36	63.2	74	12 Q9IHS8	Q9IHS8 tt virus. o

ALIGNMENTS

RESULT 1	Q8WP81	PRELIMINARY:	PRT:	533 AA.
ID	Q8WP81			
AC	Q8WP81			
DT	01-MAR-2002 (TRENBLER, 20, Created)			
DT	01-MAR-2002 (TRENBLER, 20, Last sequence update)			
DT	01-JUN-2002 (TRENBLER, 21, Last annotation update)			
DE	F13B12.6 protein.			
GN	F13B12.6.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;			
OC	Rhabditidae; Pelodierinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Bardill S.C.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MDLINF-99069613; PubMed=9851916;			
RA	none.			
RT	"genome sequence of the nematode C.elegans: A platform for			
RL	investigating biology."			
RL	Science 282:2012-2018(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Wild A.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; Z70686; CAD21656.1; JOINED.			
DR	EMBL; Z70683; CAD21626.1; JOINED.			
DR	EMBL; Z70686; CAD21626.1; JOINED.			
DR	EMBL; Z70686; CAD21626.1; JOINED.			
DR	InterPro: IPR000980; SH2.			
DR	Pfam: PF00017; SH2.1.			
DR	ProDom: PD000093; SH2.1.			
DR	SMART; SM00252; SH2.1.			
DR	PROSITE; P55001; SH2.1.			
SQ	SEQUENCE 533 AA: 60897 MW: C244CF1884D08FF4 CRC64:			

17	36	63.2	139	2 Q54919	Q54919 streptococ
18	36	63.2	187	5 Q9VW20	Q9VW20 drosophila
19	36	63.2	231	2 Q9XC33	Q9XC33 streptococ
20	36	63.2	326	5 Q62010	Q62010 ceratilis c
21	36	63.2	329	10 Q9SM15	Q9SM15 zea mays (m
22	36	63.2	482	16 Q9YXG0	Q9YXG0 brucea me
23	36	63.2	546	16 Q9A9J1	Q9A9J1 caulobacter
24	36	63.2	681	5 Q9X0S8	Q9X0S8 caenorhabdi
25	36	63.2	781	16 Q9ZK65	Q9ZK65 thizobium m
26	36	63.2	1082	11 Q55201	Q55201 mus musculu
27	36	63.2	1084	13 Q9DDT5	Q9DDT5 brachydantio
28	36	63.2	1087	4 Q43279	Q43279 homo sapien
29	36	63.2	1087	4 Q96399	Q96399 homo sapien
30	36	63.2	1087	4 Q00267	Q00267 homo sapien
31	36	63.2	2658	13 Q9DMF0	Q9DMF0 gallus gall
32	36	63.2	4060	12 Q9IHZ8	Q9IHZ8 gill-associ
33	35	61.4	115	11 Q8VHX8	Q8VHX8 mus musculu
34	35	61.4	153	12 Q70700	Q70700 human calic
35	35	61.4	201	10 Q9FE62	Q9FE62 arabidopsis
36	35	61.4	249	4 Q9H8H7	Q9H8H7 homo sapien
37	35	61.4	315	16 Q9L1T5	Q9L1T5 streptomyc
38	35	61.4	368	5 Q9V655	Q9V655 drosophila
39	35	61.4	369	5 Q9V655	Q9V655 drosophila
40	35	61.4	369	5 Q8E747	Q8E747 drosophila
41	35	61.4	420	5 Q18380	Q18380 drosophila
42	35	61.4	435	5 Q9VB08	Q9VB08 drosophila
43	35	61.4	446	5 Q9VZC2	Q9VZC2 drosophila
44	35	61.4	483	16 Q9PEB1	Q9PEB1 xyella fas

Query Match 70.2%; Score 40; DB 5; Length 533;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 SSTRAPOP 11
DB 185 SSTRAPOP 193

RESULT 2

O98U88 PRELIMINARY; PRT; 224 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MADH dehydrogenase I, C subunit.
GN DRI504.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
RN NCBI_TaxID=1299;
RP SEQUENCE FROM N.A.
RC STRAIN-R1.
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moritz K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Ueberback T., Zaleski C.,
RA Makarova K.A., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1";
RL Science 286:1571-1577(1999).
DR TYGR: A601994; AAF11071.1; .
DR InterPro: IPR001268; Complex1_30K.
DR Pfam: PF00329; Complex1_30Kd.1.
DR ProDom: PD001581; Complex1_30Kd.1.
DR PROSITE: PS00542; COMPLEX1_30K; 1.
KW Complete proteome.
SQ SEQUENCE 224 AA; 25032 MW; 05CC3BDF11281B35 CRC64;

Query Match 68.4%; Score 39; DB 16; Length 224;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 SSTRAPOP 11
DB 79 SSTRAPOP 87

RESULT 3

O98SL7 PRELIMINARY; PRT; 587 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE EYES absent-1 alpha.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
RN NCBI_TaxID=8355;
RP SEQUENCE FROM N.A.
RC STRAIN-972H-1.
RA MEDLINE=21233041; PubMed=11335132;
RA David R., Ahrens K., Wedlich D., Schlosser G.;
RT "Xenopus Eyal demarcates all neurogenic placodes as well as migrating
RT hypaxial muscle precursors";

RL Mech. Dev. 103:189-192(2001).
DR EMBL: AF352028; AAK31354.1; .
DR InterPro: IPR001454; Hlganase/hydrlase.
DR Pfam: PF00702; Hydrolase; 1.
SQ SEQUENCE 587 AA; 64142 MW; 5EE9813F6075ADE4 CRC64;

Query Match 68.4%; Score 39; DB 13; Length 587;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 SSTRAPOP 10
DB 162 SSTRAPOP 169

RESULT 4

O98SL6 PRELIMINARY; PRT; 592 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE EYES absent-1 beta.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
RN NCBI_TaxID=8355;
RP SEQUENCE FROM N.A.
RC STRAIN-972H-1.
RA MEDLINE=21233041; PubMed=11335132;
RA David R., Ahrens K., Wedlich D., Schlosser G.;
RT "Xenopus Eyal demarcates all neurogenic placodes as well as migrating
RT hypaxial muscle precursors";
RL Mech. Dev. 103:189-192(2001).
DR EMBL: AF352028; AAK31355.1; .
DR InterPro: IPR001454; Hlganase/hydrlase.
DR Pfam: PF00702; Hydrolase; 1.
SQ SEQUENCE 592 AA; 64640 MW; 84BE9396231B2C6 CRC64;

Query Match 68.4%; Score 39; DB 13; Length 592;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 SSTRAPOP 10
DB 167 SSTRAPOP 174

RESULT 5

O9P6N4 PRELIMINARY; PRT; 906 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Putative vacuolar biogenesis protein, yeast endi homolog.
GN SPAC823.12.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetaceae.
RN NCBI_TaxID=4896;
RP SEQUENCE FROM N.A.
RC STRAIN-972H-1.
RA Votkxert G., Aert R., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL355013; CAB90157.1; .
DR InterPro: IPR000547; Clathrin_repeat.
DR Pfam: PF00637; Clathrin; 1.
DR SMART: SM00299; CLH; 1.

DR SMART; SM00184; RING: 1.
SQ SEQUENCE 906 AA; 102523 MW; 74F8B5B4440F3C80 CRC64;
Query Match
Best Local Similarity 68.4%; Score 39; DB 3; Length 906;
Best Local Similarity 87.5%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 SYTAPOPO 11
|||||
Db 630 SYTAPOPO 637

RESULT 6
ID 090H57 PRELIMINARY: PRT: 124 AA.
AC 090H57; (TREMBLER. 13, Created)
DT 01-MAY-2000 (TREMBLER. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLER. 13, Last annotation update)
DE UL33 homolog.
OS Gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Infectious latrogotrichitis-like viruses.
OX NCBI_TaxID=10386;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-SA-2;
RA Johnson M.A.;
RT "Sequence of the infectious latrogotrichitis virus (SA-2 strain)
RT unique long region UL28 to UL43."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF168792; AAD56207.1; FE67191F6E3C7DA CRC64;
SQ SEQUENCE 124 AA; 14267 MW; FE67191F6E3C7DA CRC64;
Query Match
Best Local Similarity 66.7%; Score 38; DB 12; Length 124;
Best Local Similarity 54.5%; Pred. No. 12;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 QASSYTAPOPO 11
|:|:|:|
Db 4 QSNHSTPOPO 14

RESULT 7
ID 0949H3 PRELIMINARY: PRT: 295 AA.
AC 0949H3;
DT 01-DEC-2001 (TREMBLER. 19, Created)
DT 01-DEC-2001 (TREMBLER. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLER. 21, Last annotation update)
DE Putative class I chitinase (Fragment).
GN R030.
OS Hevea brasiliensis (Para rubber tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Malpighiales; Euphorbiaceae; Hevea.
OX NCBI_TaxID=3981;
RN (1)
RP SEQUENCE FROM N.A.
RC O'Riordan G.; Godric-Cavar J.; Radauer C.; Hoffmann-Sommergruber K.;
RA Scheiner O.; Epner C.; Breiteneder H.;
RT "Cloning and expression of a class I chitinase from latex, Hevea
RT brasiliensis."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ238579; GAC42881.1;
DR InterPro: IPR001002; Chitin_binding_1.
DR InterPro: IPR000726; Glyco_hydro_19.
DR Pfam: PF00187; Chitin_bind_1;
DR Pfam: PF00182; Glyco_hydro_19;
DR Pfam: PD000574; Glyco_hydro_19;
DR PROSITE: PS00773; CHITINASE_19_1; UNKNOWN_1.
DR PROSITE: PS00774; CHITINASE_19_2; UNKNOWN_1.
KW Chitin-binding.

FT NON PER 1
SQ SEQUENCE 295 AA; 31647 MW; 6B748B2B3BFCF96 CRC64;
Query Match
Best Local Similarity 66.7%; Score 38; DB 10; Length 295;
Best Local Similarity 70.0%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QASSYTAPOPO 10
|||||
Db 142 QASSYCPSP 151

RESULT 8
ID 08VDM6 PRELIMINARY: PRT: 859 AA.
AC 08VDM6;
DT 01-MAR-2002 (TREMBLER. 20, Created)
DT 01-MAR-2002 (TREMBLER. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLER. 21, Last annotation update)
DE Similar to E1B-55 kDa-associated protein 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC Strauberg R.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC021506; AAN1506.1;
DR EMBL: BC021506; AAN1506.1;
DR InterPro: IPR003034; SAP.
DR InterPro: IPR003878; SPRY domain.
DR InterPro: IPR003877; SPRY_receptor.
DR Pfam: PF02037; SAP; 1.
DR Pfam: PF00622; SPRY; 1.
DR SMART: SM00513; SAP; 1.
DR SMART: SM00449; SPRY; 1.
SQ SEQUENCE 859 AA; 96002 MW; 197328B681DF260E CRC64;
Query Match
Best Local Similarity 66.7%; Score 38; DB 11; Length 859;
Best Local Similarity 87.5%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;

OY 3 SSYTAPOPO 10
|||||
Db 751 SSYTAPOPO 758

RESULT 9
ID P78364 PRELIMINARY: PRT: 1004 AA.
AC P78364;
DT 01-MAY-1997 (TREMBLER. 03, Created)
DT 01-MAY-1997 (TREMBLER. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLER. 19, Last annotation update)
DE Polyhomeotic 1 homolog (Fragment).
GN HP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE-97220024; PubMed-9121482;
RA Gunster M.J.; Satijn D.P.; Hamer K.M.; den Blaauwen J.L.;
RT "Identification and characterization of interactions between the
RT vertebrate polycomb-group protein Bmi1 and human homologs of
RT polyhomeotic."
RL Mol. Cell. Biol. 17:2326-2335(1997).
DR EMBL: U89277; AAC51168.1;
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM; 1.
DR SMART: SM00454; SAM; 1.

FT NON-TER 1004 1004
SQ SEQUENCE 1004 AA; 105479 MW; D2A61904C7D32EB1 CRC64;
Query Match
Best Local Similarity 66.7%; Score 38; DB 4; Length 1004;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QASSTAPQ 10
DB 441 QATLTAPQ 450

RESULT 10

Q9RMZ8 PRELIMINARY; PRT; 90 AA.
AC Q9RMZ8
DT 01-MAY-2000 (Tremblrel. 13, Created)
DE 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE Hypothetical protein DR0517.
GN DR0517.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanathavan J.S., Lam P., McDonald L., Uterback T., Zaleski C.,
RA Ketchum K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE001910; AAF10097.1; -.
DR TIGR: DR0517; -.
SQ Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 8948 MW; F6DB1B9F1FDE5D CRC64;
Query Match
Best Local Similarity 64.9%; Score 37; DB 16; Length 90;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 ASSYAPQ 11
DB 36 ASSYAPQ 45

RESULT 11

Q8YBW7 PRELIMINARY; PRT; 112 AA.
AC Q8YBW7
DT 01-MAR-2002 (Tremblrel. 20, Created)
DE 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Potassium efflux system protein PHAC.
GN BM110769.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RA MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,

RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009711; AAL54011.1; -.
DR InterPro: IPR002757; DUF67.
DR Pfam: PF01898; DUF67.1.
DR ProDom: PD006097; DUF67.1.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12010 MW; D386BE18F5EE81B CRC64;

Query Match
Best Local Similarity 64.9%; Score 37; DB 16; Length 112;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 QASSTAPQ 11
DB 64 QAAQYTDYPPQ 74

RESULT 12

Q9U044 PRELIMINARY; PRT; 142 AA.
AC Q9U044
DT 01-MAY-2000 (Tremblrel. 13, Created)
DE 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 16.6 kDa protein (Fragment)
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 H90;
RA MEDLINE=20223868; PubMed=10759889;
RA Ding D.O., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hirooka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
DR EMBL: AB027821; BA87125.1; -.
DR InterPro: IPR001140; ABCtransprtm.
DR Pfam: PF00664; ABC_membrane; 1.
KW Hypothetical protein.
FT NON-TER 142 142
SQ SEQUENCE 142 AA; 16561 MW; 8296B1B8BD0B5C4F CRC64;
Query Match
Best Local Similarity 64.9%; Score 37; DB 3; Length 142;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 SSTYAPQ 11
DB 73 SSTYAPQ 81

RESULT 13

Q9XE77 PRELIMINARY; PRT; 422 AA.
AC Q9XE77
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE Patatin-like protein.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Paludoidae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;

```

RN [1]
RP SEQUENCE FROM N.A.
RA Liaca V., Lou A., Messing J.W.;
RT "Microsytent analysis of 22-kDa zein cluster in maize and sorghum.";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF061282; ADD22169.1; -
DR InterPro: IPR002641; Patatin.
DR Pfam: PF01734; Patatin; 1.
SQ SEQUENCE 422 AA; 45766 MW; 742128BBDA37F479 CRC64;

Query Match
Best Local Similarity 77.8%; Score 37; DB 10; Length 422;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 QASSTAPQ 9
Db 288 QAEKTAPO 296

RESULT 14
O61711 PRELIMINARY; PRT; 583 AA.
ID 061711;
AC 061711;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 64.6 kDa protein.
GN R119.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RC MEDLINE=99069613; PubMed=9851916;
RX None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RC Clarke K., Bauer C., O'Neal D.;
RT "The sequence of C. elegans cosmid R119.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RC Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF063007; AAC16433.1; -
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KW Hypothetical protein.
SQ SEQUENCE 583 AA; 64623 MW; B7DBE9E3DFICEF13 CRC64;

Query Match
Best Local Similarity 64.9%; Score 37; DB 5; Length 583;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 QASSTAPQ 11
Db 125 QASSTAPNP 135

RESULT 15
O9P5NO PRELIMINARY; PRT; 1465 AA.

```

```

AC O9P5NO:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Probable ATP-dependent permease c359.05.
GN SPBC359.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA McCombie W.R.;
RT "Sequence analysis of a region of the fission yeast genome.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS). MRP SUBFAMILY.
DR EMBL: AL356012; CAB91574.1; -
DR HSSP: P13569; INRD
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABC_transport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00064; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR PRODOM: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA_1
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein;
KW Transport.
RN [1]
RP TRANSMEM
FT DOMAIN 1 8
FT TRANSMEM 9 29
FT TRANSMEM 30 42
FT TRANSMEM 43 63
FT TRANSMEM 64 69
FT TRANSMEM 70 90
FT TRANSMEM 91 97
FT TRANSMEM 98 118
FT TRANSMEM 119 129
FT TRANSMEM 130 150
FT TRANSMEM 151 291
FT TRANSMEM 292 312
FT TRANSMEM 313 362
FT TRANSMEM 363 383
FT TRANSMEM 384 394
FT TRANSMEM 395 415
FT TRANSMEM 416 480
FT TRANSMEM 481 501
FT TRANSMEM 502 899
FT TRANSMEM 900 920
FT TRANSMEM 921 939
FT TRANSMEM 940 960
FT TRANSMEM 961 1033
FT TRANSMEM 1034 1054
FT TRANSMEM 1055 1133
FT TRANSMEM 1134 1154
FT TRANSMEM 1155 1465
FT NP_BIND 614 621
FT NP_BIND 1260 1267
FT NP_BIND 1268 1267
FT CARBOHYD 346 346
FT CARBOHYD 1076 1076
SQ SEQUENCE 1465 AA; 166612 MW; B15F1579B0895332 CRC64;

Query Match
Best Local Similarity 64.9%; Score 37; DB 3; Length 1465;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 SSYTAPQ 11
Db 279 SSYTSPNP 287

```

Mon Jan 13 10:35:14 2003

Search completed: January 13, 2003, 09:59:45
Job time : 44.7436 secs

us-09-554-941-2.rspt

Gencore version 5.1.3
2003 Compugen Ltd

using sw model
+time 11 Seconds

41.4/0

09-554-941-2

```

Sequence:
--ocnm62
--next 0.5

```

112892
41476328 loc-
segs,
moters:

Total number of
length: 0

Maximilian first
existing first

pred. NO. greater than or equal to the score greater by analysis of the and is derived by analysis of the

SUMMARIES

SwissProt_40:*

SUMMARY

Descriptor	Frequency	Percentage
1. <u>Very good</u>	1	100
2. <u>Good</u>	0	0
3. <u>Fair</u>	0	0
4. <u>Poor</u>	0	0
5. <u>Very poor</u>	0	0
6. <u>Not applicable</u>	0	0
7. <u>Other</u>	0	0
8. <u>Don't know</u>	0	0
9. <u>Refused</u>	0	0
10. <u>Blank</u>	0	0
11. <u>Other</u>	0	0
12. <u>Don't know</u>	0	0
13. <u>Refused</u>	0	0
14. <u>Blank</u>	0	0
15. <u>Other</u>	0	0
16. <u>Don't know</u>	0	0
17. <u>Refused</u>	0	0
18. <u>Blank</u>	0	0
19. <u>Other</u>	0	0
20. <u>Don't know</u>	0	0
21. <u>Refused</u>	0	0
22. <u>Blank</u>	0	0
23. <u>Other</u>	0	0
24. <u>Don't know</u>	0	0
25. <u>Refused</u>	0	0
26. <u>Blank</u>	0	0
27. <u>Other</u>	0	0
28. <u>Don't know</u>	0	0
29. <u>Refused</u>	0	0
30. <u>Blank</u>	0	0
31. <u>Other</u>	0	0
32. <u>Don't know</u>	0	0
33. <u>Refused</u>	0	0
34. <u>Blank</u>	0	0
35. <u>Other</u>	0	0
36. <u>Don't know</u>	0	0
37. <u>Refused</u>	0	0
38. <u>Blank</u>	0	0
39. <u>Other</u>	0	0
40. <u>Don't know</u>	0	0
41. <u>Refused</u>	0	0
42. <u>Blank</u>	0	0
43. <u>Other</u>	0	0
44. <u>Don't know</u>	0	0
45. <u>Refused</u>	0	0
46. <u>Blank</u>	0	0
47. <u>Other</u>	0	0
48. <u>Don't know</u>	0	0
49. <u>Refused</u>	0	0
50. <u>Blank</u>	0	0
51. <u>Other</u>	0	0
52. <u>Don't know</u>	0	0
53. <u>Refused</u>	0	0
54. <u>Blank</u>	0	0
55. <u>Other</u>	0	0
56. <u>Don't know</u>	0	0
57. <u>Refused</u>	0	0
58. <u>Blank</u>	0	0
59. <u>Other</u>	0	0
60. <u>Don't know</u>	0	0
61. <u>Refused</u>	0	0
62. <u>Blank</u>	0	0
63. <u>Other</u>	0	0
64. <u>Don't know</u>	0	0
65. <u>Refused</u>	0	0
66. <u>Blank</u>	0	0
67. <u>Other</u>	0	0
68. <u>Don't know</u>	0	0
69. <u>Refused</u>	0	0
70. <u>Blank</u>	0	0
71. <u>Other</u>	0	0
72. <u>Don't know</u>	0	0
73. <u>Refused</u>	0	0
74. <u>Blank</u>	0	0
75. <u>Other</u>	0	0
76. <u>Don't know</u>	0	0
77. <u>Refused</u>	0	0
78. <u>Blank</u>	0	0
79. <u>Other</u>	0	0
80. <u>Don't know</u>	0	0
81. <u>Refused</u>	0	0
82. <u>Blank</u>	0	0
83. <u>Other</u>	0	0
84. <u>Don't know</u>	0	0
85. <u>Refused</u>	0	0
86. <u>Blank</u>	0	0
87. <u>Other</u>	0	0
88. <u>Don't know</u>	0	0
89. <u>Refused</u>	0	0
90. <u>Blank</u>	0	0
91. <u>Other</u>	0	0
92. <u>Don't know</u>	0	0
93. <u>Refused</u>	0	0
94. <u>Blank</u>	0	0
95. <u>Other</u>	0	0
96. <u>Don't know</u>	0	0
97. <u>Refused</u>	0	0
98. <u>Blank</u>	0	0
99. <u>Other</u>	0	0
100. <u>Don't know</u>	0	0
101. <u>Refused</u>	0	0
102. <u>Blank</u>	0	0
103. <u>Other</u>	0	0
104. <u>Don't know</u>	0	0
105. <u>Refused</u>	0	0
106. <u>Blank</u>	0	0
107. <u>Other</u>	0	0
108. <u>Don't know</u>	0	0
109. <u>Refused</u>	0	0
110. <u>Blank</u>	0	0
11		

RESULT 1
GMOA.DACDE STANDARD: PRT: 680 AA.

ID GMOA.DACDE
AD 001745; 1993 (Rel. 26, Created)
IC 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 40, Last annotation update) (GMO).
DT 16-OCT-2001 (Rel. 40, Last annotation update) (GMO).
DE galactose oxidase precursor (EC 1.1.3.9) (GMO).
DE Galactose oxidase (Cladobotryum dendroides).
OS Dactylium dendroides (Cladobotryum dendroides).
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes.
OC Hypocreales; Hypocreales; Hypomyces.
NCBI_taxid=5133;
[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
RA MEDLINE=92235025; Pubmed=1569070; yadav K.D.S., Keen J.N.,
RA McPherson M.J., Ogel Z.B., Stevens C.E., yadav K.D.S., Keen J.N.,
RA Knowles P.F.;
RA "galactose oxidase of Dactylium dendroides. Gene cloning and sequence
analysis". Chem. 267:8146-8152(1992).
RA j. Biol. Chem. 267:8146-8152(1992).
RL [2] X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RA MEDLINE=91163341; Pubmed=2002850; Ogel Z.B., McPherson M.J.,
RA Ito N., yadav K.D.S., Knowles P.F.;
RA Keen J.N., yadav K.D.S., revealed by a 1.7 A crystal structure
RA "Novel thioether bond revealed by a 1.7 A crystal structure of
galactose 3,5O:8T-90(1991)."
RU FUNCTION: CATALYSES THE STEROSPECIFIC OXIDATION OF A BROAD RANGE
OF PRIMARY ALCOHOL SUBSTRATES.
CC - FUNCTION: CATALYSES THE STEROSPECIFIC OXIDATION OF A BROAD RANGE
OF PRIMARY ACTIVITY: D-galactose + O(2) = D-galacto-hexodiallose +
CC H(2O)(2).
CC -1 CATALYST: COPPER.
CC H(2O)(2).
CC -1 SUBUNIT: MONOMER.
CC -1 SUBCELLULAR LOCATION: 3 KELCH REPEATS.
CC -1 SIMILARITY: CONTRAINS 3 KELCH REPEATS.
CC -1 SIMILARITY: NAME=wothington-biochem.com/manual/g/GMO.html.
CC -1 DATA=http://www.worthington-biochem.com/manual/g/GMO.html.
CC WWW=
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation at
the European Bioinformatics Institute. There are no restrictions on any
use by non-profit institutions as long as its content and for commercial
use by non-profit institutions is not removed. See http://www.isb.ch/announce/
CC modified requires a license@isp-sib.ch).
CC or send an email to license@isp-sib.ch).
CC CC
CC CC
CC EMBL; 150F; 31-JAN-94.
DR PDB; 150G; 31-JAN-94.
DR PDB; 150H; 31-JAN-94.
DR PDB; 150I; 31-JAN-94.
DR InterPro; IPRO0198; Kelch.
DR Pfam; PF00501; KELCHREPEAT.
DR PRINTS; PR00501; KELCHREPEAT.

Q06916 myxococcus
P19198 dictyostell
Q35188 mus musculu
Q9FG24 arabidopsis
Q9FW68 delnoccus
P49541 odonellia s
Q17711 caenorhabd
P46589 candida alb
P20156 rattus norv
Q67708 aquifex neo
P49115 cavia porce
P35831 mus musculu

DR SMART: SM00231; PAS8C; 1.
KW OXIDOREDUCTASE; Signal: Copper; Repeat: 3d-structure.
FT CHAIN 1 41
FT REPEAT 276 680
FT REPEAT 323 321
FT REPEAT 492 372
FT DISULFID 59 544
FT DISULFID 269 68
FT METAL 556 313
FT METAL 537 313
FT METAL 536 313
FT METAL 537 313
FT METAL 622 622
FT STRAND 48 49
FT STRAND 56 59
FT STRAND 63 63
FT STRAND 64 65
FT STRAND 66 66
FT HELIX 68 72
FT STRAND 76 77
FT STRAND 80 81
FT STRAND 85 87
FT STRAND 94 112
FT STRAND 120 123
FT STRAND 123 121
FT STRAND 125 131
FT STRAND 142 145
FT STRAND 153 169
FT STRAND 172 172
FT STRAND 173 174
FT STRAND 180 181
FT STRAND 183 188
FT STRAND 197 198
FT STRAND 201 207
FT STRAND 214 217
FT STRAND 219 221
FT STRAND 224 227
FT STRAND 236 237
FT STRAND 243 249
FT STRAND 250 253
FT STRAND 254 260
FT STRAND 262 263
FT STRAND 271 274
FT STRAND 276 277
FT STRAND 279 283
FT STRAND 285 286
FT STRAND 288 289
FT STRAND 291 295
FT STRAND 296 299
FT STRAND 300 303
FT STRAND 312 312
FT STRAND 315 318
FT STRAND 320 321
FT STRAND 324 327
FT STRAND 330 330
FT STRAND 332 332
FT STRAND 340 344
FT STRAND 345 348
FT STRAND 349 353
FT STRAND 354 355
FT STRAND 359 361
FT HELIX 366 368
FT STRAND 369 372
FT STRAND 373 373
FT STRAND 378 380
FT STRAND 382 384
FT STRAND 386 388
FT STRAND 394 399
FT STRAND 402 403
FT STRAND 405 411
FT STRAND 413 414
FT STRAND 415 416

us-09-554-941-2.rsp

FT STRAND 417 418
FT STRAND 423 423
FT STRAND 424 424
FT STRAND 426 431
FT STRAND 432 432
FT STRAND 435 435
FT STRAND 436 440
FT STRAND 443 443
FT STRAND 450 450
FT STRAND 452 452
FT STRAND 455 459
FT STRAND 463 464
FT STRAND 468 471
FT STRAND 473 474
FT STRAND 480 481
FT STRAND 484 487
FT STRAND 489 490
FT STRAND 493 496
FT STRAND 499 499
FT STRAND 501 501
FT STRAND 504 505
FT STRAND 510 510
FT STRAND 515 518
FT STRAND 519 522
FT STRAND 523 526
FT STRAND 536 537
FT STRAND 538 542
FT STRAND 544 545
FT STRAND 548 552
FT STRAND 557 558
FT STRAND 565 570
FT STRAND 572 574
FT STRAND 575 575
FT STRAND 577 578
FT STRAND 581 581
FT STRAND 587 590
FT STRAND 594 596
FT STRAND 597 598
FT STRAND 600 605
FT STRAND 611 616
FT STRAND 619 621
FT STRAND 622 623
FT STRAND 624 624
FT STRAND 630 633
FT STRAND 636 637
FT STRAND 642 646
FT STRAND 651 653
FT STRAND 657 664
FT STRAND 666 667
FT STRAND 670 670
FT STRAND 674 679
SQ SEQUENCE 680 AA; 72823 MM; 2P97C561B63E46E9 CMC64;
Query Match 91.28; Score 52; DB 1; Length 680;
Best Local Similarity 100.08; Pred No. 0.059; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QASSY7APOP 10
DB 188 QASSY7APOP 197
RESULT 2
WDRC_HUMAN
AC Q9GZL7; Q9NV18; Q96HU0; Q9NV80; PRT; 423 AA.
DT 15-JUN-2002 (Rel. 41, Created)
DR 15-JUN-2002 (Rel. 41, Last sequence update)
DB WD-repeat protein 12 (YTM1 homology).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

NCBI_TaxID=9606;
 NCBL [1]
 SEQUENCE FROM N.A.
 RA Matsumoto S.;
 RA "Human homolog of Saccharomyces cerevisiae YTM1."
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiraori A., Sudo H., Sugawara M.,
 RA Wagaatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Takiguchi S.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Oho Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuko Y.,
 RA Niimiya K., Iwayanagi T.;
 RA "NDO human cDNA sequencing project."
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC Tissue-Skin;
 RA Strausberg R.;
 RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF242546; AAF60355.1;
 DR EMBL: AK001743; BAA91875.1;
 DR EMBL: AK022781; BAB14242.1;
 DR EMBL: AK022782; BAB14243.1;
 DR EMBL: BC008082; AAH08082.1;
 DR Genbank: HENC114098; WDR12.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINRPT.
 DR ProDom: PD000018; WD40; 2.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS00678; WD_REPEATS_1; 2.
 DR PROSITE: PS00682; WD_REPEATS_2; 5.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat; Polymorphism.
 FT REPEAT 99 137 WD 1.
 FT REPEAT 137 180 WD 2.
 FT REPEAT 180 226 WD 3.
 FT REPEAT 226 253 WD 4.
 FT REPEAT 253 293 WD 5.
 FT REPEAT 293 334 WD 6.
 FT REPEAT 334 380 WD 7.
 FT REPEAT 380 422 WD 7.
 FT REPEAT 422 72 M -> V.
 FT VARIANT 72 72 /FTId=VAR_012863.
 FT VARIANT 89 89 Y -> C.
 FT VARIANT 89 89 /FTId=VAR_012864.
 FT VARIANT 286 286 E -> G.
 FT VARIANT 286 286 /FTId=VAR_012865.
 FT CONFLICT 75 75 I -> V (IN REF. 2; BAA91875).
 FT CONFLICT 333 333 L -> M (IN REF. 1).
 FT SEQUENCE 423 AA; 47708 MW; 6D088C640AC981D8 CRC64;
 SQ
 Query Match 64.9%; Score 37; DB 1; Length 423;
 Best Local Similarity 85.7%; Pred No. 19;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
 WDR12_MOUSE STANDARD: PRT; 423 AA.
 ID WDR12_MOUSE
 AC 09J04; Q9J0F5; Q9J0F3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WD-repeat protein 12 (YTM1 homolog).
 GN WDR12
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Thymus;
 RA MEDLINE=2168157; PubMed=11827460;
 RA Nal B., Mohr E., Da Silva M.I., Tagett R., Navarro C., Carroll P.,
 RA Deperis D., Verthuy C., Jordan B.R., Ferrier P.;
 RA "Wdr12, a mouse gene encoding a novel WD-repeat protein with a
 RA notchless-like amino-terminal domain."
 RA Genomics 79:77-86(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ X C57BL/6J;
 RA Matsumoto S.;
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; Tissue-Brain;
 RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
 RA Hashimoto K.;
 RA "Isolation of full-length cDNA clones from mouse brain cDNA library
 RA made by oligo-capping method."
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue-Testis, and Pancreas;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Arakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischer W., Gaasterland T., Gissi C., King B., Kochava H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli I., Sakamoto N.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Storch K.-F.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch L.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RA Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC Strausberg R.;
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

```

RESULT 4
ID STRN_STRGR STANDARD; PRT; 384 AA.
AC P03399;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE dtdp-dihydrostreptose-streptidine-6-phosphate
DE dihydrostreptosyltransferase (EC 2.4.2.27).
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RX MEDLINE=91375432; PubMed=1654502;
RA Mansouri K., Plepersberg W.;
RT "Genetics of streptomycin production in Streptomyces griseus:
RT nucleotide sequence of five genes, strFGH1, including a phosphatase
RT gene."
RL Mol. Gen. Genet. 228:459-469(1991).
CC -1- FUNCTION. PROBABLY IS A DIHYDROSTREPTOSYL GLYCOSYLTRANSFERASE,
CC INVOLVED IN THE FIRST GLYCOSYLATION STEP CONDENSING STREPTIDINE-
CC 6-PHOSPHATE AND DIHYDROSTREPTOSE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: dtdp-L-dihydrostreptose + streptidine 6-
CC phosphate = dtdp + O-1,4-alpha-L-dihydrostreptosyl-streptidine 6-
CC phosphate.
CC -1- PATHWAY: Streptomycin biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its

```

RP SEQUENCE FROM N.A.
RA MEDLINE-21638749; Pubmed-11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavarides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow R.F., Bates K.N., Beard L.M., Beate D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Gardner C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cooley V.E., Collier R.E., Connor R.E., Corry N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliland R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jevonch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Jevonch K., Johnson C.M., Johnson D.,
RA Levansthalo M.H., Levensha M., King A., Knights A., Lloyd G., Lawlor S.,
RA Marsh V.L., Martin S.L., McConnechie L.J., Cloyd D.M., Lowell J.D.,
RA Milne S.A., Mistry D., Moore M.J.F., Mulligan J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Senha H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Sodeltund C., Stewart C.A., Sulston J.E.,
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RT Rogers J.;
RT The DNA sequence and comparative analysis of human chromosome 20.;
RL Nature 414:865-871(2001).
RP [3]

RX MEDLINE=95359195; PubMed=7632726;
 RA Chen K.-S., Deluca H.F.;
 RT "Cloning of the human I alpha,25-dihydroxyvitamin D-3 24-hydroxylase
 RT gene promoter and identification of two vitamin D-responsive
 RT elements.";
 RT Biochim. Biophys. Acta 1263:1-9(1995).
 RN [4]
 RP SEQUENCE OF 258-514 FROM N.A.
 RX MEDLINE=94091187; PubMed=826631;
 RA Labuda M., Lemieux N., Tily F., Prinster C., Giorleux F.H.;
 RT "Human 25-hydroxyvitamin D 24-hydroxylase cytochrome P450 subunit
 RT maps to a different chromosomal location than that of pseudovitamin
 RT D-deficient rickets.";
 RL J. Bone Miner. Res. 8:1397-1406(1993).
 CC -1- FUNCTION: HAS A ROLE IN MAINTAINING CALCIUM HOMEOSTASIS. CATALYZES
 CC THE NADPH-DEPENDENT 24-HYDROXYLATION OF 25-HYDROXYVITAMIN D(3) IN
 CC THE PRESENCE OF ADRENODOXIN AND NADPH-ADRENODOXIN REDUCTASE.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: L13286; AAA62379.1; -;
 DR EMBL: AL138805; CAB91829.1; -;
 DR EMBL: U60669; AAB03776.1; ALT_SEQ.
 DR EMBL: S67623; AAB29308.1; -;
 DR HSP: P14779; JLP2.
 DR Genew; HGNC:2602; CYP24.
 DR MIM: 126065; -;
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Heme;
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 35 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 36 514 CYTOCHROME P450-CC24.
 FT BINDING 462 462 HEME (POTENTIAL).
 FT CONFLICT 68 68 G -> A (IN REF. 1).
 FT CONFLICT 124 125 AY -> V (IN REF. 1).
 FT CONFLICT 270 270 D -> G (IN REF. 1).
 FT CONFLICT 365 365 V -> R (IN REF. 1).
 FT CONFLICT 368 368 A -> E (IN REF. 1).
 FT CONFLICT 390 390 S -> G (IN REF. 1).
 FT CONFLICT 511 511 F -> S (IN REF. 4).
 SQ SEQUENCE 514 AA; 58875 MW; 8662F63771981195 CRC64;
 Query Match 63.2%; Score 36; DB 1; Length 514;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ASSYTAPOPO 11
 DB 30 STAYTSPQPR 39
 RESULT 6
 TRA_BPMU STANDARD; PRT; 663 AA.
 AC P07636; P06021;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transposase.
 GN A OR 3.
 OS Bacteriophage Mu.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC Mu-like viruses.
 OX NCBI_TaxID=10677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86067968; PubMed=2999776;
 RA Harshey R.M., Getzoff E.D., Baldwin D.L., Miller J.L., Chaconas G.;
 RT "Primary structure of phage mu transposase: homology to mu
 RT repressor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7676-7680(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Priess H., Brauer B., Schmidt C., Kamp D.;
 RT "Sequence of the left end of Mu.";
 RL (in) Symonds N., Roussaint A., Van de Putte P., Howe M.M. (eds.);
 RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,
 RL New York (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Morgan G., Hattall G., Hendrix R.;
 RT "Genome of bacteriophage Mu and comparison with the Haemophilus
 RT influenzae Mu-like prophage Flumu.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-88 FROM N.A.
 RX MEDLINE=83012203; PubMed=6214696;
 RA Priess H., Kamp D., Kahmann R., Brauer B., Delius H.;
 RT "Nucleotide sequence of the immunity region of bacteriophage Mu.";
 RL Mol. Gen. Genet. 186:315-321(1982).
 RN [5]
 RP SEQUENCE OF 1-84 FROM N.A.
 RX MEDLINE=83218562; PubMed=6222246;
 RA Toussaint A., Faelen M., Desmet L., Allet B.;
 RT "The products of gene A of the related phages Mu and D108 differ in
 RT their specificities.";
 RL Mol. Gen. Genet. 190:70-79(1983).
 RN [6]
 RP STRUCTURE BY NMR OF 1-76.
 RX MEDLINE=95187707; PubMed=7881904;
 RA Clubb R.T., Omichinski J.G., Savillanti H., Mizuuchi K.,
 RA Gronenborn A.M., Clore G.M.;
 RT "A novel class of winged helix-turn-helix protein: the DNA-binding
 RT domain of Mu transposase.";
 RL Structure 2:1041-1048(1994).
 RN [7]
 RP STRUCTURE BY NMR OF 76-174.
 RX MEDLINE=98070329; PubMed=9405381;
 RA Schumacher S., Clubb R.T., Cai M., Mizuuchi K., Clore G.M.,
 RA Gronenborn A.M.;
 RT "Solution structure of the Mu end DNA-binding Ibeta subdomain of
 RT phage Mu transposase: modular DNA recognition by two tethered
 RT domains.";
 RL EMBO J. 16:7532-7541(1997).
 RN [8]
 RP STRUCTURE BY NMR OF 173-247.
 RX MEDLINE=98035037; PubMed=9367742;
 RA Clubb R.T., Schumacher S., Mizuuchi K., Gronenborn A.M., Clore G.M.;
 RT "Solution structure of the I gamma subdomain of the Mu end
 RT DNA-binding domain of phage Mu transposase.";
 RL J. Mol. Biol. 273:19-25(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 248-574.
 RX MEDLINE=95354202; PubMed=7628012;
 RA Rice P., Mizuuchi K.;
 RT "Structure of the bacteriophage Mu transposase core: a common
 RT structural motif for DNA transposition and retroviral integration.";
 RL Cell 82:209-220(1995).
 CC -1- FUNCTION: THIS TRANSPOSASE IS ESSENTIAL FOR INTEGRATION,
 CC REPLICATION-TRANSPOSITION, AND EXCISION OF MU DNA.
 CC -1- MISCELLANEOUS: MU CAN TRANSPOSE ITS DNA INTO MULTIPLE SITES IN
 CC MANY BACTERIAL GENOMES AND MEDIATE A VARIETY OF DNA
 CC REARRANGEMENTS. TRANSPOSITION REQUIRES BOTH TRANSPOSASE (ENCODED
 CC BY GENE A) AND TRANSPOSITION ENHANCER (ENCODED BY GENE B).
 CC -1- MISCELLANEOUS: UNLIKE OTHER TRANSPOSONS MU HAS DISSIMILAR

CC SEQUENCES AT ITS LEFT AND RIGHT ENDS. TRANSPOSASE APPARENTLY BINDS
 CC 3 SPECIFIC BLOCKS OF SEQUENCES AT EACH END OF MU DNA.
 CC -1- MISCELLANEOUS: THE A GENE IS REGULATED BY THE REPRESSOR C, WHICH
 CC BINDS TO AN OPERATOR SEQUENCE & TURNS OFF TRANSCRIPTION. REPRESSOR
 CC ON CAN, AT HIGH CONCENTRATIONS, OCCUPY ALMOST THE EXACT SAME SITES
 CC ON MU ENDS AS THE TRANSPOSASE, AND TRANSPOSASE CAN BIND TO
 CC FRAGMENTS CONTAINING THE MU OPERATOR SEQUENCE.
 CC -1- SIMILARITY: STRONG, TO H. INFLUENZAE H11478.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M11195; AAA2369.1; -
 CC EMBL: M64097; AAA2379.1; -
 CC EMBL: AF083977; AAF01083.1; -
 CC EMBL: V01464; CAA24713.1; -
 CC EMBL: V00868; CAA24236.1; -
 CC PIR: A24746; TOBPU.
 CC DR PIR: 14-FEB-95.
 CC DR PDB: 1TNT; 14-FEB-95.
 CC DR PDB: 1BCM; 15-OCT-95.
 CC DR PDB: 1BCO; 15-OCT-95.
 CC DR PDB: 2EZH; 03-DEC-97.
 CC DR PDB: 2E2I; 03-DEC-97.
 CC DR PDB: 2E2K; 14-JAN-98.
 CC DR PDB: 2E2L; 14-JAN-98.
 CC DR InterPro: IPR003314; Mu_DNA_bind.
 CC DR InterPro: IPR004189; Mu_transposase.
 CC DR Pfam: PF02316; Mu_DNA_bind.1;
 CC DR Pfam: PF02914; Mu_transposase.1;
 CC KM Transposition; Transposable element; DNA-binding; DNA excision;
 CC KW DNA integration; DNA recombination; 3D-structure.
 CC FT DNA_BIND 35 55 H-T-H MOTIF (POTENTIAL).
 CC FT DNA_BIND 390 409 H-T-H MOTIF (POTENTIAL).
 CC FT CONFLICT 66 66 G -> R (IN REF. 5).
 CC FT CONFLICT 408 408 P -> S (IN REF. 2).
 CC SQ SEQUENCE 663 AA; 75003 MW; B882CFDCBFC0B2E3 CRC64;
 CC
 CC Query Match 63.2%; Score 36; DB 1; Length 663;
 CC Best Local Similarity 60.0%; Pred. No. 46;
 CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 ASSTAPPOQ 11
 CC DB 411 AGAYTGPMPQ 420
 CC
 CC RESULT 7
 CC CASB_CAMDR STANDARD; PRT; 232 AA.
 CC AC Q9TYD0;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DE 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Beta casein precursor.
 CC GN CSN2.
 CC OS Camelus dromedarius (Dromedary) (Arabian camel).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
 CC NCBI_TaxID=9838;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Somali; TISSUE=Mammary gland;
 CC RX MEDLINE=98291310; PubMed=9627840;
 CC RA Kappellet S., Farah Z., Puhon Z.,
 CC RA "Sequence analysis of Camelus dromedarius milk caseins".
 CC RA J. Dairy Res. 65:209-222(1998).
 CC RL -1- FUNCTION: IMPORTANT ROLE IN DETERMINATION OF THE SURFACE

CC PROPERTIES OF THE CASEIN MICELLES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CASEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ012630; CAA10079.1; -
 CC DR InterPro: IPR001588; Casein.
 CC DR Pfam: PF00363; caseins.1.
 CC DR PROSITE: PS00306; CASEIN_ALPHA-BETA; FALSE_NEG.
 CC KW Milk; Phosphorylation; Glycoprotein; Signal.
 CC FT SIGNAL 1 15 BY SIMILARITY.
 CC FT CHAIN 16 232 BETA CASEIN.
 CC FT MOD_RES 30 30 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 33 33 PHOSPHORYLATION (BY SIMILARITY).
 CC FT MOD_RES 34 34 PHOSPHORYLATION (BY SIMILARITY).
 CC SQ SEQUENCE 232 AA; 26218 MW; A0F9F41D2EA7C518 CRC64;
 CC
 CC Query Match 61.4%; Score 35; DB 1; Length 232;
 CC Best Local Similarity 63.6%; Pred. No. 23;
 CC Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 QASSTAPPOQ 11
 CC DB 62 QDKITFPPOQ 72
 CC
 CC RESULT 8
 CC IF2M_SCHPO STANDARD; PRT; 686 AA.
 CC ID IF2M_SCHPO
 CC AC 059683;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Translation initiation factor IF-2, mitochondrial precursor (IF-2Mt)
 CC DE (IF-2Mt).
 CC GN SPBC2F6.01 OR SPBC1271.15C.
 CC OS Schizosaccharomyces pombe (Fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21840401; PubMed=11859360;
 CC RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 CC RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 CC RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 CC RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 CC RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 CC RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 CC RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 CC RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 CC RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 CC RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,
 CC RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
 CC RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesti D., Hilbert H.,
 CC RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 CC RA Eger P., Zimmermann W., Wedler H., Wambutt R., Parnelle B.,
 CC RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Motiller S.,
 CC RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Spharovsky G.V., Useary D., Barrell B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*,"
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION
 CC OF PROTEIN SYNTHESIS. PROTECTS FORMYL METHIONYL-tRNA FROM
 CC SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S
 CC RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP
 CC DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AL034353; CNA22205.1;
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR00178; IF2.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00009; GTP_EFTU_1.
 DR Pfam: PF03144; GTP_EFTU_D2; 2.
 DR Prodom: PD186100; IF2; 1.
 DR Prodom: PD1860231; small_GTP; 1.
 DR TIGRFAMs: TIGR00487; IF-2; 1.
 DR TIGRFAMs: TIGR00487; IF-2; 1.
 DR PROSITE: PS01176; IF2; 1.
 DR Initiation factor; protein biosynthesis; GTP-binding; Transit peptide;
 KM Mitochondrion.
 KW TRANSIT ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 686 TRANSLOCATION INITIATION FACTOR IF-2.
 FT DOMAIN 170 322 G-DOMAIN.
 FT NP_BIND 178 185 GTP (BY SIMILARITY).
 FT NP_BIND 226 230 GTP (BY SIMILARITY).
 FT NP_BIND 280 283 GTP (BY SIMILARITY).
 FT SEQUENCE 686 AA; 75611 MW; 6F5B5CDECF3A5CD CRC64;
 SQ
 Query Match 61.4%; Score 35; DB 1; Length 686;
 Best Local Similarity 45.5%; Pred. No. 73;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 OY 1 QASSTAPPOQ 11
 DB 148 EAAKSPSPK 158
 RESULT 9
 FLNA_HUMAN STANDARD; PRT; 2647 AA.
 ID FLNA_HUMAN
 AC P21333;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Filamin A (Alpha-filamin) (Filamin 1) (Endothelial actin-binding
 DE protein) (ABP-280) (Nonmuscle filamin).
 GN FLNA OR FLN1 OR FLN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP MEDLINE=90361737; PubMed=2391361;
 RX Gorlin J.B., Yamin R., Egan S., Stewart M., Stossel T.P.,
 RA Kwiatkowski D.J., Hartwig J.H.,
 RA "Human endothelial actin-binding protein (ABP-280, nonmuscle

RT filamin): a molecular leaf spring.";
 RL J. Cell Biol. 111:1089-1105(1990).
 RL [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=96311563; PubMed=8733135;
 RA Chen E.Y., Zollo M., Mazzarella R.A., Ciocciolella A., Chen C.-N.,
 RA Zhu L., Heiner C., Burrough F.W., Ripetto M., Schlusinger D.,
 RA D'Urso M.;
 RT "Long-range sequence analysis in Xq28: thirteen known and six
 RT candidate genes in 219.4 Kb of high GC DNA between the RCP/GCP and
 RT G6PD loci.";
 RL Hum. Mol. Genet. 5:659-668(1996).
 RL [3]
 RP SEQUENCE OF 1658-1772 FROM N.A.
 RP MEDLINE=93357748; PubMed=7689010;
 RA Meistrini E., Patrosso C., Mancini M., Rivella S., Rocchi M.,
 RA Repetto M., Villa A., Fratini A., Zoppe M., Vezoni P.,
 RA Tonello D.;
 RT "Mapping of two genes encoding isoforms of the actin binding protein
 RT ABP-280, a dystrophin like protein, to Xq28 and to chromosome 7.";
 RL Hum. Mol. Genet. 2:761-766(1993).
 RL [4]
 RP REVIEW.
 RP MEDLINE=21234905; PubMed=11336782;
 RA van der Flier A., Sonnenberg A.;
 RT "Structural and functional aspects of filamins.";
 RL Biochim. Biophys. Acta 1538:99-117(2001).
 RL [5]
 RP REVIEW.
 RP MEDLINE=21146932; PubMed=11252955;
 RA Stossel T.P., Condellis J., Cooley L., Hartwig J.H., Noegel A.,
 RA Schleicher M., Shapiro S.S.;
 RT "Filamins as integrators of cell mechanics and signalling.";
 RL Nat. Rev. Mol. Cell Biol. 2:138-145(2001).
 RL [6]
 RP VARIANTS PH PHE-656 AND THR-1764.
 RP MEDLINE=21423782; PubMed=11532987;
 RA Shuen V.L., Dixon P.H., Fox J.W., Hong S.E., Kinton L., Sisodia S.M.,
 RA Duncan J.S., Dubois F., Scheffer I.E., Schachter S.C., Wilner A.,
 RA Henchy R., Crino P., Kanuro K., Dimario F., Berg W., Kuzniack R.,
 RA Cole A.J., Bromfield E., Biber M., Schomer D., Whelless J., Silver R.,
 RA Mochida G.H., Berkovic S.F., Andermann F., Andermann E., Dobyns W.B.,
 RA Wood N.W., Walsh C.A.;
 RT "Mutations in the X-linked filamin 1 gene cause periventricular
 RT nodular heterotopia in males as well as in females.";
 RL Hum. Mol. Genet. 10:1775-1783(2001).
 CC -1- FUNCTION: Promotes orthogonal branching of actin filaments and
 CC links actin filaments to membrane glycoproteins. Anchors various
 CC transmembrane proteins to the actin cytoskeleton and serves as a
 CC scaffold for a wide range of cytoplasmic signaling proteins.
 CC -1- SUBUNIT: Homodimer. Interacts with cyHSP. Interacts with various
 CC other binding partners. PERIPHERAL CYTOPLASM.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL CYTOPLASM.
 CC -1- TISSUE SPECIFICITY: UNIDIRECTIONAL.
 CC -1- PTM: PHOSPHORYLATION EXTENT CHANGES IN RESPONSE TO CELL
 CC ACTIVATION.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: Defects in FLNA are the cause of periventricular
 CC heterotopia (PH) also called nodular heterotopia, bilateral
 CC periventricular (NHP or BPNH). PH is an X-linked developmental
 CC dominant disorder in which many neurons fail to migrate into the
 CC cerebral cortex. They remain as nodules lining the ventricular
 CC surface. In heterozygous females these neurons presumably
 CC contain the active X chromosome with the filamin mutation. Most
 CC hemizygous affected males die early during embryogenesis, whereas
 CC heterozygous females have normal intelligence but suffer from
 CC seizures and various manifestations outside the central nervous
 CC system, especially related to the vascular system. This implies
 CC that essential embryonic cell migration can only occur in FLNA-
 CC expressing cells.
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,

CC ABP-120, ABP-180, OR BETA-FODRIN).
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FILAMIN REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X53416; CAA37495.1; -
 DR EMBL: L44140; AAA92844.1; -
 DR EMBL: X70082; CAA49687.1; -
 DR EMBL: X70085; CAA49690.1; -
 DR PIR: A37098; A37098.
 DR HSSP: P13466; 1KSR.
 DR GeneW: HGNC:3754; FLNA.
 DR MIM: 300017; -
 DR MIM: 300049; -
 DR InterPro: IPR001589; Actbind_actuin.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR001298; Filamin.
 DR Pfam: PF00307; CH; 2.
 DR Pfam: PF00630; Filamin; 24.
 DR SMART: SM00033; CH; 2.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00020; ACTININ_2; 1.
 DR PROSITE: PS50194; FILAMIN_REPEAT; 24.
 DR Actin-binding: Phosphorylation; Repeat; Polymorphism;
 KW Disease mutation.
 FT DOMAIN 1 274
 FT 1 149
 FT DOMAIN 1 149
 FT 1 149
 FT REPEAT 166 266
 FT 276 374
 FT REPEAT 376 474
 FT 475 570
 FT REPEAT 571 663
 FT 663 763
 FT REPEAT 764 866
 FT 866 965
 FT REPEAT 966 1061
 FT 1062 1154
 FT REPEAT 1155 1249
 FT 1250 1349
 FT REPEAT 1350 1442
 FT 1443 1539
 FT REPEAT 1540 1636
 FT 1636 1740
 FT REPEAT 1741 1778
 FT 1779 1860
 FT REPEAT 1861 1950
 FT 1951 2039
 FT REPEAT 2042 2131
 FT 2132 2230
 FT REPEAT 2233 2325
 FT 2327 2420
 FT REPEAT 2424 2516
 FT 2517 2551
 FT REPEAT 2552 2646
 FT 2647 2662
 FT DOMAIN 1761 1762
 FT 1762 1762
 FT REPEAT 320 320
 FT 320 320
 FT REPEAT 370 370
 FT 370 370
 FT REPEAT 552 552
 FT 552 552
 FT REPEAT 656 656
 FT 656 656
 FT REPEAT 1764 1764
 FT 1764 1764

FT FT
 FT CONFLICT 1772 1772 /FTId-VAR.012835.
 FT FT 2634 2634 A -> G (IN REF. 3).
 SO SEQUENCE 2647 AA; 280759 MW; 6C1A07041DFA3D42 CRC64;
 Query Match 61.48; Score 35; DB 1; Length 2647;
 Best Local Similarity 100.08; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 YTAAP 10
 DB 1712 YTAAP 1717

RESULT 10
 ID PKAA_STRCO STANDARD; PRT; 543 AA.
 AC P54739;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-OCT-1996 (Rel. 34, Last sequence update)
 DE Serine/threonine protein kinase Pkaa (EC 2.7.1.-).
 GN Pkaa OR SCO2974 OR SCE50.02C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 ON NCBI_TaxID=1902;
 RP [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN-A3(2);
 RX MEDLINE=9518913; PubMed=7883195;
 RA Urahe H., Ogawara H.;
 RT "Cloning, sequencing and expression of serine/threonine kinase-
 RL gene 153.99-104(1995)."
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Krieser H.,
 RA Cronin A., Fraser A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Huang C.-H., Krieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2)."
 RT Nature 417:141-147(2002).
 RL -1- PM: AUTOPHOSPHORYLATED MAINLY AT THR AND SLIGHTLY AT SER.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D86821; BA013168.1; -
 DR EMBL: AL163672; CAB87324.1; -
 DR HSSP: O63450; 1A06.
 DR PhosSite: P54739; -
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_Thr_Pkinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Complete proteome.
 FT DOMAIN 8 276 PROTEIN_KINASE.
 FT NP_BIND 14 22 ATP (BY SIMILARITY).
 FT BINDING 48 48 ATP (BY SIMILARITY).
 FT ACT_SITE 142 142 BY SIMILARITY.
 FT DOMAIN 301 481 GLN/PRO-RICH
 SO SEQUENCE 543 AA; 58181 MM; 0E1965520FA0C200 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 543;
 Best Local Similarity 54.5%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 QASSTAPQPO 11
 DB 437 QPQRYATPPO 447

RESULT 11
 VE4_HPV54 STANDARD; PRT; 134 AA.
 ID 081022:
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Probable E4 protein.
 GN E4.
 OS Human papillomavirus type 54.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=37113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delius H.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U37488; AA79191.1;
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 KW Early protein.
 SO SEQUENCE 134 AA; 15373 MM; B157979B6035EE2A CRC64;

Query Match 57.9%; Score 33; DB 1; Length 134;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QASSTAPQ 9
 DB 96 QTSNHTAPQ 104

RESULT 12
 CPC3_CANPG STANDARD; PRT; 175 AA.
 ID P81584;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cuticle protein CP1876 (CPCP1876).
 OS Cancer pagurus (Rock crab).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Decapoda; Pleocyemata;
 OC Brachyura; Pterychotura; Cancroidea; Cancridae; Cancer.
 OX NCBI_TaxID=6755;

RN [1]
 RP SEQUENCE.
 RC Tissue-cuticle;
 RX MEDLINE=99354472; PubMed=10425740;
 RA Andersen S.O.;
 RT "Exoskeletal proteins from the crab, Cancer pagurus.";
 RL Comp. Biochem. Physiol. 123A:203-211(1999).
 CC -1- TISSUE SPECIFICITY: CALCIFIED SHELL.
 CC -1- MASS SPECTROMETRY: MW=18759.8; METHOD=MALDI.
 KW Structural protein; Cuticle.
 SO SEQUENCE 175 AA; 18765 MM; 57C1C83D75CB132E CRC64;

Query Match 57.9%; Score 33; DB 1; Length 175;
 Best Local Similarity 60.0%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 ASSYAPQPO 11
 DB 43 ADVYMPRPQ 52

RESULT 13
 FL3L_HUMAN STANDARD; PRT; 235 AA.
 ID P49771;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3
 DE ligand) (Flt3L).
 GN FLT3LG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94195428; PubMed=8145851;
 RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
 RA Bazan J.F., Kastelein R., Hudak S., Wagner J., Mattson J., Luth J.,
 RA Duda G., Martini N., Peterson D., Menon S., Shanaleit A.,
 RA Zelenik A., Rosner O., Dubreuil P., Barnbaum D., Lee F.;
 RA "Cloning of the human homologue of the murine flt3 ligand: a growth
 RA ligand for FLT3/Flt3 receptor tyrosine kinase regulates growth of
 RA hematopoietic stem cells and is encoded by variant RNAs.";
 RL Nature 368:643-648(1994).
 RL
 RL
 RL
 RP SEQUENCE FROM N.A.
 RP MEDLINE=94235842; PubMed=8180375;
 RA Lyman S.D., James L., Johnson L., Brasel K., de Vries P.,
 RA Escobar S.S., Downey H., Splet R.R., Beckmann M.P., McKenna H.J.;
 RA "Cloning of the human homologue of the murine flt3 ligand: a growth
 RA factor for early hematopoietic progenitor cells.";
 RT Blood 83:2795-2801(1994).
 RL
 RL
 RL
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RP MEDLINE=96032581; PubMed=7566977;
 RA Lyman S.D., Stocking K., Davison B., Fletcher F., Johnson L.,
 RA Escobar S.;
 RA "Structural analysis of human and murine flt3 ligand genomic loci.";
 RT Oncogene 11:1165-1172(1995).
 RL
 RL
 RL
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE=20343011; PubMed=10881197;
 RA Savvides S.N., Boone T., Karplus P.A.;
 RA "Flt3 ligand structure and unexpected commonalities of helical
 RA bundles and cystine knots.";
 RT Nat. Struct. Biol. 7:486-491(2000).
 CC -1- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC
 CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING
 CC FACTORS AND INTERLEUKINS.
 CC -1- SUBUNIT: Homodimer (isoform 2).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);

CC secreted (isoform 2).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a 1/membrane-bound (shown here)
CC and 2/soluble; are produced by alternative splicing.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U04806; AAA17999.1; -
DR EMBL: U03858; AAA19825.1; -
DR EMBL: U29874; AAA90949.1; -
DR EMBL: U29874; AAA90950.1; -
DR PDB: 1ETE; 09-JUN-00.
DR Genbank: HGNC:3766; FLT3LG.
DR MIM: 600007; -
DR InterPro: IPR004213; FLT3_1lg.
DR Pfam: PF02947; FLT3_1lg; 1.
KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal;
KW 3D-structure.
FT SIGINL 1 26
FT CHAIN 27 235
FT DOMAIN 27 184
FT TRANSMEM 185 205
FT DOMAIN 206 235
FT DISULFID 30 111
FT DISULFID 70 153
FT DISULFID 119 158
FT CARBOHD 126 126
FT CARBOHD 149 149
FT VANSPLC 161 178
FT VANSPLC 179 235
FT CONFLICT 72
SQ SEQUENCE 235 AA; 26416 MW; 738958F693B4CECF CRC64;
Query Match
Best Local Similarity 57.9%; Score 33; DB 1; Length 235;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QASSTAPAP 10
Db 175 EATAPAP 184
ID OMEY_CHLPPN STANDARD: PRT; 262 AA.
AC 0926M5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative outer membrane protein CPN1034/CPN1034 precursor.
GN CPN1034 OR CP0818 OR CPJ1034
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CM1029;
RA MEDLINE=99206506; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RT Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RL "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.",
Natl. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RA MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Eisen J., Nelson M., Deboy R., Kolony J., McClarty G., Salzberg S.L.,
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT pneumoniae AR39.",
RN Nucleic Acids Res. 28:1397-1406(2000).
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RA MEDLINE=2030349; PubMed=10871362;
RA Shira M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.,
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.",
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001684; AAD19171.1; -
DR EMBL: AE002240; AAR3613.1; -
DR EMBL: AP002548; BAA99241.1; -
DR TIGR: CP0818; -
KW Hypothetical protein; Outer membrane; Signal; Complete proteome.
FT CHAIN 1 17
FT SIGNAL 18 262
SQ SEQUENCE 262 AA; 30000 MW; 22D30D12FEC5A893 CRC64;
Query Match
Best Local Similarity 57.9%; Score 33; DB 1; Length 262;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TAPOPQ 11
Db 148 TAPOPQ 153
ID GDA4_WHEAT STANDARD: PRT; 297 AA.
AC P04724;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alpha/beta-gliadin A-IV precursor (Prolamin).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Scrophophyta; Embryophyta; Tracheophyta;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85234522; PubMed=2989281;
RA Okita T.W., Cheesbrough V., Reeves C.D.,
RT "Evolution and heterogeneity of the alpha-/beta-type and gamma-type
RT gliadin DNA sequences.",
RL J. Biol. Chem. 260:8203-8213(1985).
CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
CC HOMOLOGUE CLASSES: THE ALPHA/BETA-GLIADINS CAN BE DIVIDED INTO 5
CC TO SINGLE BASE SUBSTITUTIONS AND TO DUPLICATIONS OR DELETIONS
CC WITHIN OR NEAR DIRECT REPEATS. THERE ARE MORE THAN A 100 COPIES OF
CC THE GENE FOR ALPHA/BETA-GLIADIN PER HAPLOID GENOME.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M11075; AAA34282.1; -
 DR PIR; D22364; D22364.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001954; Gila_glutenh.
 DR InterPro; IPR001768; Try/amy1_inhbr.
 DR Pfam; PF00234; ctyp_alpha_amy1; 1.
 DR PRINTS; PR00208; GLIADGLOTFN.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Repeat; Signal; Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 297 ALPHA/BETA-GLIADIN A-IV.
 SQ SEQUENCE 297 AA; 34239 MW; 0025ED289AE9588B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 297;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 STAPPOQ 11
 II IIII
 DB 106 STPOPOQ 113

Search completed: January 13, 2003, 09:54:07
 Job time : 13.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2003, 09:53:20 ; Search time 10.1538 Seconds
(without alignments)
104.146 Million cell updates/sec

Title: US-09-554-941-2
Perfect score: 57
Sequence: 1 QASSTAPQPQ 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	91.2	728	1 A38084	galactose oxidase
2	40	70.2	1416	2 T20823	hypothetical prote
3	39	68.4	224	2 B75388	NADH dehydrogenase
4	37	64.9	90	2 G75509	hypothetical prote
5	37	64.9	112	2 A43605	potassium efflux s
6	37	64.9	583	2 G87719	protein R119.7 (lm
7	36	63.2	139	2 S31941	penicillin-binding
8	36	63.2	384	2 S17778	strH protein - str
9	36	63.2	482	2 AE3381	ATP-dependent RNA
10	36	63.2	513	1 A47436	1,25-dihydroxyvita
11	36	63.2	546	2 A87372	xylian 1,4-beta-xy
12	36	63.2	662	1 A87372	transposase - phag
13	36	63.2	681	2 T23455	hypothetical prote
14	36	63.2	1082	2 T42204	chromatin structur
15	35	61.4	483	2 B82722	hypothetical prote
16	35	61.4	686	2 T33951	translation initia
17	35	61.4	820	2 T46412	ubiquitin-protein
18	35	61.4	874	2 T34922	phosphoenolpyruvat
19	35	61.4	875	1 A57080	cell surface antis
20	35	61.4	1124	2 T30340	dsRNA adenosine de
21	35	61.4	2129	2 T27431	hypothetical prote
22	35	61.4	2647	2 A37098	gelation factor AB
23	34	59.6	151	2 T08561	hypothetical prote
24	34	59.6	336	2 E84594	AP2 domain transcr
25	34	59.6	373	2 G72368	single-stranded DN
26	34	59.6	440	2 G84631	hypothetical prote
27	34	59.6	453	2 F84560	purple acid phosph
28	34	59.6	453	2 T17823	hypothetical cytos
29	34	59.6	466	1 T04599	acid phosphatase (

30	34	59.6	487	2 H82298	D-alanyl-D-alanine
31	34	59.6	502	2 T08776	hypothetical prote
32	34	59.6	525	1 A75570	2-oxo acid dehydro
33	34	59.6	543	1 A75570	protein kinase (EC
34	34	59.6	616	2 E84424	proline auxin tra
35	34	59.6	640	2 G96733	auxin transport pr
36	34	59.6	706	2 G71004	hypothetical prote
37	34	59.6	818	1 J43397	peroxinectin precu
38	34	59.6	856	2 T13159	E1B-55kDa-associat
39	34	59.6	1012	2 T43162	RAE-28 - mouse
40	34	59.6	1747	2 T43162	vitellogenin - gyp
41	33	57.9	64	2 AD3228	isopentenyladenosine
42	33	57.9	182	2 G70687	hypothetical prote
43	33	57.9	205	2 AH3407	hypothetical cytos
44	33	57.9	235	2 I38440	fl3 ligand - huma
45	33	57.9	251	2 T14456	MADS box protein h

ALIGNMENTS

RESULT 1
A38084
galactose oxidase (EC 1.1.3.9) precursor [validated] - fungus (Cladobotryum dendroides)
C:Species: Cladobotryum dendroides
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #extl_change 20-Oct-2000
C:Accession: A38084
R:McPherson, M.J.; Ogel, Z.B.; Stevens, C.; Yadav, K.D.S.; Keen, J.N.; Knowles, P.F.
J. Biol. Chem. 267, 8146-8152, 1992
A:Title: Galactose oxidase of Dactylium dendroides. Gene cloning and sequence analysis.
A:Reference number: A38084; MUID:92235025; PMID:1569070
A:Accession: A38084
A:Molecule type: DNA
A:Residues: 1-728 (MCP>
A:Cross-references: GB:M86819
A:Note: It is uncertain whether Met-1 or Met-49 is the initiator
A:Note: parts of this sequence, including the amino end of the mature protein, were c
R:Itto, N.; Phillips, S.E.V.; Stevens, C.; Ogel, Z.B.; McPherson, M.J.; Keen, J.N.; Y
Nature 350, 87-90, 1991
A:Title: Novel thioether bond revealed by a 1.7 angstrom crystal structure of galacto
A:Reference number: A48244; MUID:91163641; PMID:2002850
A:Contents: annotation
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51740; PDB:1GOF
A:Contents: annotation, X-ray crystallography, 1.7 angstroms, residues 90-728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51741; PDB:1GOG
A:Contents: annotation, X-ray crystallography, 1.9 angstroms, residues 90-728
R:Itto, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A:Reference number: A51742; PDB:1GOF
A:Contents: annotation, X-ray crystallography, 2.2 angstroms, residues 90-728
C:Genetics:
A:Gene: gaoc
A:Superfamily: Cladobotryum dendroides galactose oxidase
A:Keywords: disulfide bond; metal binding; oxidoreductase
F:1-64/Domain: signal sequence #status predicted <SIG>
F:55-89/Domain: propeptide #status predicted <PRO>
F:90-728/Product: galactose oxidase #status experimental <MAT>
F:107-116,604-607/Disulfide bonds: #status experimental
F:317-361/Cross-link: cysteinyltyrosine (Cys-Tyr) #status experimental
F:361,584,585,670/Binding site: copper (Tyr, Tyr, His, His) #status experimental
F:419/Binding site: substrate (Arg) #status predicted

Query Match 91.2% ; Score 52; DB 1; Length 728;
Best Local Similarity 100.0% ; Pred. No. 0.091;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QASSTAPQP 10
|||||||
DB 236 QASSTAPQP 245

RESULT 2

T20823
 hypothetical protein R10H10.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T20823; T24156
 R:Wild, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19329
 A:Accession: T20823
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1416 <W12>
 A:Cross-references: EMBL:Z70683; PDB:CAA94594.1; GSPDB:GN00022; CESP:R10H10.7
 R:Barclay, S.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19846
 A:Accession: T24156
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1416 <W12>
 A:Cross-references: EMBL:Z70686; PDB:CAA94614.1; GSPDB:GN00022; CESP:R10H10.7
 C:Genetics:
 A:Gene: CESP:R10H10.7
 A:Map position: 4
 A:Insertions: 83/3; 254/1; 355/2; 472/3; 548/1; 600/3; 650/1; 733/3; 843/2; 976/3; 1080/3;
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein R10H10.7

Query Match

Best Local Similarity 70.2%; Score 40; DB 2; Length 1416;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 SRYTAPQ 11
 |||:||||

Db 185 SRYTAPQ 193

RESULT 3

B75388
 NADH dehydrogenase I, C subunit - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75388
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: B75388
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <WMT>
 A:Cross-references: GB:AE001994; GB:AE000513; NID:96459259; PDB:AAFI1071.1; PID:9645926
 C:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1504
 A:Map position: 1

Query Match

Best Local Similarity 68.4%; Score 39; DB 2; Length 224;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 SRYTAPQ 11
 |||:||||

Db 79 SRYTAPQ 87

RESULT 4

G75509

hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75509
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75509
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-90 <WMT>
 A:Cross-references: GB:AE001910; GB:AE000513; NID:96458198; PDB:AAFI0097.1; PID:9645
 C:Genetics:
 A:Gene: DR0517
 A:Map position: 1

Query Match

Best Local Similarity 64.9%; Score 37; DB 2; Length 90;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 ASSYAPQ 11
 |||:||||

Db 36 ASSYAPQ 45

RESULT 5

AH3605
 potassium efflux system protein phac [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 C:Accession: AH3605
 R:DeIvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AH3605
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-112 <KUR>
 A:Cross-references: GB:AE008918; PDB:AL54011.1; PID:917984962; GSPDB:GN00191
 C:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI10769
 A:Map position: 11
 C:Superfamily: *Pyrococcus abyssi* hypothetical protein PAB1887

Query Match

Best Local Similarity 64.9%; Score 37; DB 2; Length 112;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 QASSYAPQ 11
 ||:||||

Db 64 QAAQYDPVQ 74

RESULT 6

G87719
 protein R119.7 [imported] - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: G87719
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

us-09-554-941-2.rpt

Breaches
 Matches
 1 QMSSTAPPOQ 11
 11111 1135
 QY 125 QMSSTAPPNPR 135
 DB
 Streptococcus pneumoniae (strain 100511) (fragment
 range 26-May-2000
 41Ks

```

penicillin-binding protein IX  

Streptococcus pneumoniae  

EMBL accession: F06789.1  

Strain: 100511  

Date: 06-Jan-1994  

Author(s): Jordens, J.; Daniels, M.; Coffey, T.J.; Bates, J.; Paul, C.  

Source: The EMBL Data Library, February 1993  

Molecule type: DNA  

Accession number: S31941  

Description: Penicillin-resistant pneumococci isolated in  

resistance type: 9 <KEY>  

Accession: 100511  

Database: EMBL  

Cross-referenced source: penicillin-binding protein 1B  

Experimental protocol: penicillin resistance  

Keywords: antibiotic resistance  

Query Match Similarity 63.2%; Score 36; DB 2; Length 139; Gaps 0;  

Best Local 6; Conservative 6; 7%; Pred. No. 16; Indels 0; Mismatches 3;
Matches 3
SSTYAPPOQ 11
1:::|||||
OY 84 SSTYAPPOQ 92
DB *text_change 22-Oct-1999

```

[illegible]

```

RESULT 9
NEE381      ATP-dependent RNA helicase dead BME11035 [imported] - Brucella melitensis
AE38381     ATP-dependent Brucella melitensis sequence, revision 01-Feb-2002 #text-change 01-Feb-2002
C1:Species: 01-Feb-2381
C1:Date:    01-Feb-2381
C1:Author:  W.G.; Kapralov, V.; Redkar, R.J.; Patra, G.; Majer, C.; Ios, T.; Ivanov
C1:Acc:Vecchio, G.; Goldsman, U.S.A. of the facultative intracellular pathogen Brucella melitensis
R1:Del:Kazur, M. Acad. Sci. sequence PMID:11756688
..Proc: Natl. Gen. genome AD3252; PMID:11756688
..PID: 917982999; GSPDB:GN00190
A:File: The number 361;
A:Reference: AE3381;
A:Accession: Prelim.DNA
A:Stable type: 1.482 <KUR>
A:Motif:ques: 1.482 <KUR>
A:Res:references: GB: strain 10M
A:Experimental source:
A:Genetics:
A:Gene: BME11035
A:Map position:
A:Map position:
Match: 63.2%; DB 2; Length 482; 0; Gaps 0;
Mismatch: 60.0%; Ered. NO. 2; Indexes
Mismatch: 2; Mismatches

```

```
Matches
1 QASQYTAPOP 10
  :|: ||| |
18 EACGYATPPT 27
Db
--base (EC 1.14.--) precursor - human
--text_change 21-Jul-2006
```

[illegible]

```

      11 - Caulobacter crescentus
      *text_change 10-May-2001
      30 SNAVTSPOPR
      2 ASSVTAPOPO 11
      matches
      0Y
      Db

```

RESULT 11
A87172, 4-beta-xylosidase (importeur)
Xylanase, 4-lobacter (presensu)
C:Accession: A87372 #sequence=evolution 20-Apr-2001
C:Accession: W.C.: Faldjyvm, T.V.: Paulsen, I.T.: Nelson, K.F.: Eisen, H.
R:Nierman, M.T.: Debroy, R.I., Dodson, R.J., Durkin, A.S.: Gann, M.L.
B.: Lapid, M.T.:

Mon Jan 13 10:35:14 2003

us-09-554-941-2.fpx

P. J. Ermolaeva, M. J. White, O. J. Saldberg, S. L. Shapiro, L. J. Venter, J. C. Fraser, C. M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4186-4191, 2001
A:Title: Complete genome sequence of *Yersinia enterocolitica* strain 480/98
A:Accession: AF073722
A:Status: Preliminary
A:Residues: 15346 <STO>
A:Cross-references: GB:AE005673; NID:913422273; PIDN:AAK22973.1; GSPDB:GN00148
C:Superfamily: Bacillus xylan 1,4-beta-xylosidase xymB

Query Match
Best Local Similarity 63.2%; Score 36; DB 1; Length 546;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 448 QASSTAPOP 10
1 QASSTAPOP 10
1 QASSTAPOP 10
1 QASSTAPOP 10

RESULT 12
Transposase - phage Mu
C:Species: Phage Mu
C:Date: 31-Mar-1988
C:Accession: A24746
R:Harshay, R. M.; G. S. 509550, 506975, 507318
A:Title: P. Acad. Sci. U.S.A. 82, 7676-7680, 1985
A:Reference: binary structure of phage mu transposase
A:Accession: A24746
A:Residues: 1662 <STO>
A:Cross-references: EMBL:86067968; PMID:2999776
R:Mu. J. 2; Characteristics: GB:M11195
A:Title: A novel 3843, 1995
A:Reference: number: 557318; MUID:95369255; PMID:7641701
A:Contents: annotated DNA binding and nuclease activity in domain III of mu transposase.
A:Title: Genet. 186, 315-321, 1992
A:Reference: nucleotide sequence of the immunity region of bacteriophage Mu.
A:Accession: number: 507291; MUID:83012203; PMID:6214696
A:Residues: 148 <STO>
A:Cross-references: EMBL:V01464; NID:915807; PIDN:CA24713.1; PID:915810
C:Function: A
C:Description: it is essential for integration, replication-transposition, and excision
C:Superfamily: Phage Mu transposase
C:Keywords: DNA binding; DNA replication

Query Match
Best Local Similarity 63.2%; Score 36; DB 1; Length 662;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 410 AOMTGPPO 11
2 AOMTGPPO 11
2 AOMTGPPO 11
2 AOMTGPPO 11

RESULT 13
Hypothetical protein K08E3.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999
R:McMurry, A.
A:Accession: T23435
A:Status: Preliminary
A:Residues: 15346 <STO>
A:Cross-references: GB:AE005673; NID:913422273; PIDN:AAK22973.1; GSPDB:GN00148
C:Superfamily: Bacillus xylan 1,4-beta-xylosidase xymB

submitted to the EMBL Data Library, November 1996

Page

A:Reference number: 219743
A:Accession: T23435
A:Status: Preliminary
A:Residues: 15346 <STO>
A:Cross-references: EMBL:86067968; PMID:2999776
R:Mu. J. 2; Characteristics: GB:M11195
A:Title: A novel 3843, 1995
A:Reference: number: 557318; MUID:95369255; PMID:7641701
A:Contents: annotated DNA binding and nuclease activity in domain III of mu transposase.
A:Title: Genet. 186, 315-321, 1992
A:Reference: nucleotide sequence of the immunity region of bacteriophage Mu.
A:Accession: number: 507291; MUID:83012203; PMID:6214696
A:Residues: 148 <STO>
A:Cross-references: EMBL:V01464; NID:915807; PIDN:CA24713.1; PID:915810
C:Function: A
C:Description: it is essential for integration, replication-transposition, and excision
C:Superfamily: Phage Mu transposase
C:Keywords: DNA binding; DNA replication

Query Match
Best Local Similarity 63.2%; Score 36; DB 1; Length 1082;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 883 QASSTAPOP 11
1 QASSTAPOP 11
1 QASSTAPOP 11
1 QASSTAPOP 11

RESULT 14
Chromatin structural protein homolog Sup5hp - mouse
C:Species: *Mus musculus* (house mouse)
C:Date: 03-Dec-1999
C:Accession: A24746
R:Harshay, R. M.; G. S. 509550, 506975, 507318
A:Title: P. Acad. Sci. U.S.A. 82, 7676-7680, 1985
A:Reference: binary structure of phage mu transposase
A:Accession: A24746
A:Residues: 1662 <STO>
A:Cross-references: EMBL:86067968; PMID:2999776
R:Mu. J. 2; Characteristics: GB:M11195
A:Title: A novel 3843, 1995
A:Reference: number: 557318; MUID:95369255; PMID:7641701
A:Contents: annotated DNA binding and nuclease activity in domain III of mu transposase.
A:Title: Genet. 186, 315-321, 1992
A:Reference: nucleotide sequence of the immunity region of bacteriophage Mu.
A:Accession: number: 507291; MUID:83012203; PMID:6214696
A:Residues: 148 <STO>
A:Cross-references: EMBL:V01464; NID:915807; PIDN:CA24713.1; PID:915810
C:Function: A
C:Description: it is essential for integration, replication-transposition, and excision
C:Superfamily: Phage Mu transposase
C:Keywords: DNA binding; DNA replication

Query Match
Best Local Similarity 63.2%; Score 36; DB 1; Length 1082;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 883 QASSTAPOP 11
1 QASSTAPOP 11
1 QASSTAPOP 11
1 QASSTAPOP 11

RESULT 15
Hypothetical protein XE1117 [Imported] - *Xylella fastidiosa* (strain 945C)
C:Species: *Xylella fastidiosa*
C:Date: 18-Aug-2000
R:Anonymous
A:Title: The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
A:Accession: B82722
A:Status: Preliminary
A:Residues: 15346 <STO>
A:Cross-references: GB:AE003947; GB:AE003849; NID:99106667; PIDN:AAF83927.1; GSPDB:GN00148
C:Superfamily: Bacillus xylan 1,4-beta-xylosidase xymB

submitted to the EMBL Data Library, November 1996

us-09-554-941-2.fpx

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
 Aulhairs, da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A: Reference number: A59328
 A: Contents: annotation
 C: Genetics:
 A: Gene: XF117

Query Match 61.4%; Score 35; DB 2; Length 483;
 Best Local Similarity 63.6%; Pred. NO. 92;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QASSTAPQPQ 11
 11: 11111
 Db 157 QATGATADAPQ 167

Search completed: January 13, 2003, 09:57:09
 Job time : 13.1538 secs

